

Tue Sep 9 08:54:24 2003

us-09-991-809-1.rapm

Davis, D.  
09/09/809 Page 1  
Seq. ID # 12/1  
11458

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2003, 16:05:03 ; Search time 171 Seconds  
(without alignments)  
63.854 Million cell updates/sec

Title: US-09-991-809-1  
Perfect score: 56  
Sequence: 1 KALVQOMEQLRQ 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues  
Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*  
32: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	12	25	US-09-991-809-1
					Sequence 1, Appli

2	56	100.0	181	37	US-60-192-739-2977	Sequence 2977, Ap
3	56	100.0	181	32	US-60-194-243-2132	Sequence 2132, Ap
4	56	100.0	244	22	US-09-791-537-140717	Sequence 140717,
5	56	100.0	377	6	US-08-256-766-2	Sequence 2, Appli
6	56	100.0	396	22	US-09-791-537-139157	Sequence 139157,
7	56	100.0	396	22	US-09-791-537-139859	Sequence 139859,
8	56	100.0	396	23	US-09-800-729-207	Sequence 207, App
9	56	100.0	396	25	US-09-987-107-33	Sequence 33, Appl
10	56	100.0	396	28	US-10-219-051B-6184	Sequence 6184, Ap
11	56	100.0	396	30	US-10-458-714-207	Sequence 207, App
12	56	100.0	396	32	US-60-452-680-24231	Sequence 24231, A
13	56	100.0	396	32	US-60-453-050-15113	Sequence 15113, A
14	56	100.0	396	32	US-60-453-135-15113	Sequence 15113, A
15	56	100.0	397	32	US-60-466-412-15113	Sequence 15113, A
16	56	100.0	397	14	US-09-079-030-123	Sequence 123, App
17	56	100.0	401	22	US-09-791-537-84940	Sequence 84940, A
18	56	100.0	401	25	US-09-987-107-36	Sequence 36, Appl
19	56	100.0	426	24	US-09-949-016-10854	Sequence 10854, A
20	56	100.0	429	22	US-09-791-537-101059	Sequence 101059,
21	56	100.0	429	25	US-09-987-107-34	Sequence 34, Appl
22	51	91.1	28	1	PCT-US95-03660-10	Sequence 10, Appl
23	51	91.1	28	6	US-08-216-537-10	Sequence 10, Appl
24	51	91.1	28	8	US-08-408-858-10	Sequence 10, Appl
25	51	91.1	376	22	US-09-791-537-139950	Sequence 139950,
26	51	91.1	396	1	PCT-US01-10670-3	Sequence 3, Appli
27	51	91.1	396	22	US-09-791-537-142101	Sequence 142101,
28	51	91.1	396	23	US-09-802-640-16	Sequence 16, Appl
29	51	91.1	396	30	US-10-403-902A-16	Sequence 16, Appl
30	49	87.5	391	22	US-09-791-537-140603	Sequence 140603,
31	49	87.5	391	23	US-09-800-729-208	Sequence 208, App
32	49	87.5	391	30	US-10-458-714-208	Sequence 208, App
33	49	87.5	395	22	US-09-791-537-142580	Sequence 142580,
34	49	87.5	395	25	US-09-987-107-35	Sequence 35, Appl
35	49	87.5	399	22	US-09-791-537-139606	Sequence 139606,
36	49	87.5	435	22	US-09-791-537-140880	Sequence 140880,
37	47	83.9	382	22	US-09-791-537-142695	Sequence 142695,
38	47	83.9	382	23	US-09-800-729-206	Sequence 206, App
39	47	83.9	382	25	US-09-987-107-37	Sequence 37, Appl
40	47	83.9	382	30	US-10-458-714-206	Sequence 206, App
41	46	82.1	10	1	PCT-US02-31642-25	Sequence 25, Appl
42	46	82.1	10	22	US-09-791-537-561	Sequence 561, App
43	46	82.1	10	22	US-09-791-537-561	Sequence 561, App
44	46	82.1	10	22	US-09-791-389-226	Sequence 226, App
45	46	82.1	10	22	US-09-791-393-226	Sequence 226, App

## ALIGNMENTS

RESULT 1  
US-09-991-809-1  
; Sequence 1, Application US/09991809  
; GENERAL INFORMATION:  
; APPLICANT: Jackowski, George  
; TITLE OF INVENTION: Apolipoprotein Biopolymer Markers Predictive of Type II Diabetes  
; FILE REFERENCE: 2132.111  
; CURRENT APPLICATION NUMBER: US/09/991.809  
; CURRENT FILING DATE: 2001-11-23  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-991-809-1

Query Match 100.0%; Score 56; DB 25; Length 12;  
Best local Similarity 100.0%; Pred. No. 0.0065;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cy 1 KALVQOMEQLRQ 12  
Db 1 KALVQOMEQLRQ 12

RESULT 2  
US-60-192-739-2977

Query Match	56;	Score	56;	DB	32;	length	181;
Best Local Similarity	100.0%;	Pred. No.	0.18;				
Matches	12;	Conservative	0;	Mismatches	0;	Indels	0;
						Gaps	0

```

RESULT 3
US-60-194-243-2132
: Sequence 2132, Application US/60194243
: GENERAL INFORMATION:
: APPLICANT: Bonazzi, Vivien
: TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: C1000417
: CURRENT APPLICATION NUMBER: US/60/194,243
: CURRENT FILING DATE: 2000-04-03
: NUMBER OF SEQ ID NOS: 3242
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2132
: LENGTH: 181
: TYPE: PRT
: ORGANISM: HUMAN
: US-60-194-243-2132

```

```
Query Match      100.0%; Score 56; DB 32; Length 181;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

```

; RESULT 4
; US-09-791-537-140717
; Sequence 140717, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danner, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 140717

```

```

; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-140717

```

Query Match	100.0%;	Score 56;	DB 22;	Length 244;
Best Local Similarity	100.0%;	Pred. No. 0.26;		
Matches 12;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

1      RESULT 5
2      US-08-256-766-2
3      Sequence 2, Application US/08256766
4      GENERAL INFORMATION:
5      APPLICANT: Deneffe, Patrice
6      APPLICANT: Guinet, Francoise
7      APPLICANT: Latina, Martine
8      APPLICANT: Murry-Brelier, Anne
9      TITLE OF INVENTION: POLYPEPTIDES DERIV
10     TITLE OF INVENTION: APOLIPOPROTEIN, F
11     NUMBER OF SEQUENCES: 26
12     CORRESPONDENCE ADDRESS:
13     ADDRESSEE: Rhone-Poulenc Rorer Inc
14     STREET: 500 Arcola Road, 3043
15     CITY: Collegeville
16     STATE: PA
17     COUNTRY: USA
18     ZIP: 19426
19     COMPUTER READABLE FORM:
20     MEDIUM TYPE: Floppy disk
21     COMPUTER: Macintosh
22     OPERATING SYSTEM: System 7.1
23     SOFTWARE: Word 5.0 (Patentin)
24     CURRENT APPLICATION DATA:
25     APPLICATION NUMBER: US/08/256,766
26     FILING DATE:
27     CLASSIFICATION: 530
28     PRIOR APPLICATION DATA:
29     APPLICATION NUMBER: FR 92/00806
30     FILING DATE: 27-JAN-1992
31     ATTORNEY/AGENT INFORMATION:
32     NAME: Goodman, Rosanne
33     REGISTRATION NUMBER: 32,534
34     REFERENCE/DOCKET NUMBER: ST92003-US
35     TELECOMMUNICATION INFORMATION:
36     TELEPHONE: (610) 454-3817
37     TELEFAX: (610) 454-3808
38     INFORMATION FOR SEQ ID NO: 2:
39     SEQUENCE CHARACTERISTICS:
40     LENGTH: 377 amino acids
41     TYPE: amino acid
42     TOPOLOGY: linear
43     MOLECULE TYPE: protein
44     US-08-256-766-2

```

Query Match	100.0%;	Score 56;	DB 6;	Length 377;
Best Local Similarity	100.0%;	Pred. No. 0.43;		
Matches 12;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

RESULT 6  
US-09-791-537-139157  
; Sequence 139157, Application US/0979153  
; GENERAL INFORMATION:  
; APPLICANT: Biomimix, Inc.

```

; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 139157
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-139157

Query Match
Best Local Similarity 100.0%; Score 56; DB 22; Length 396;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12
Db 316 KALVOQMEQLRQ 327

RESULT 7
US-09-791-139859
; Sequence 139859, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 139859
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-139859

Query Match
Best Local Similarity 100.0%; Score 56; DB 22; Length 396;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12
Db 316 KALVOQMEQLRQ 327

RESULT 8
US-09-800-729-207
; Sequence 207, Application US/09800729
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 207
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-207
```

```

Query Match
Best Local Similarity 100.0%; Score 56; DB 23; Length 396;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12
Db 316 KALVOQMEQLRQ 327

RESULT 9
US-09-987-107-33
; Sequence 33, Application US/09987107
; GENERAL INFORMATION:
; APPLICANT: GRAVERSEN, Jonas
; APPLICANT: MOESTRUP, Soren
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
; FILE REFERENCE: GRAVERSENIA
; CURRENT APPLICATION NUMBER: US/09/987,107
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/264,022
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: DK PA2001 00057
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: DK PA2000 01682
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-107-33

Query Match
Best Local Similarity 100.0%; Score 56; DB 25; Length 396;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12
Db 316 KALVOQMEQLRQ 327

RESULT 10
US-10-219-051B-6184
; Sequence 6184, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/219,051B
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 60/312,147
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/346,382
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/333,347
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 14715
; SOFTWARE: Perl script
; SEQ ID NO 6184
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: SWISS-Prot / P06727
; DATABASE ENTRY DATE: 2002-06-15
US-10-219-051B-6184

Query Match
Best Local Similarity 100.0%; Score 56; DB 28; Length 396;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 KALVOQMEQLRQ 12  
|||  
Db 316 KALVOQMEQLRQ 327

RESULT 11  
US-10-458-714-207  
; Sequence 207, Application US/10458714  
; GENERAL INFORMATION:  
; APPLICANT: NI et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1  
; CURRENT APPLICATION NUMBER: US/10/458,714  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: US/09/800,729  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 207  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-458-714-207

Query Match 100.0%; Score 56; DB 30; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.46;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KALVOQMEQLRQ 12  
|||  
Db 316 KALVOQMEQLRQ 327

RESULT 12  
US-60-452-680-24231  
; Sequence 24231, Application US/60452680  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: GRUPE, Andrew  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001450  
; CURRENT APPLICATION NUMBER: US/60/452,680  
; CURRENT FILING DATE: 2003-03-07  
; NUMBER OF SEQ ID NOS: 116213  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24231  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-452-680-24231

Query Match 100.0%; Score 56; DB 32; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.46;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KALVOQMEQLRQ 12  
|||  
Db 316 KALVOQMEQLRQ 327

RESULT 13  
US-60-453-050-15113  
; Sequence 15113, Application US/60453050  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: LUKE, May  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001457  
; CURRENT APPLICATION NUMBER: US/60/453,050  
; CURRENT FILING DATE: 2003-03-10  
; NUMBER OF SEQ ID NOS: 82762  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15113  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-453-050-15113

Query Match 100.0%; Score 56; DB 32; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.46;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KALVOQMEQLRQ 12  
|||  
Db 316 KALVOQMEQLRQ 327

RESULT 14  
US-60-453-135-15113  
; Sequence 15113, Application US/60453135  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: IAKUBOVA, Olga  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001456  
; CURRENT APPLICATION NUMBER: US/60/453,135  
; CURRENT FILING DATE: 2003-03-10  
; NUMBER OF SEQ ID NOS: 82762  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15113  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-453-135-15113

Query Match 100.0%; Score 56; DB 32; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.46;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KALVOQMEQLRQ 12  
|||  
Db 316 KALVOQMEQLRQ 327

RESULT 15  
US-60-466-412-15113  
; Sequence 15113, Application US/60466412  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: IAKUBOVA, Olga  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001466  
; CURRENT APPLICATION NUMBER: US/60/466,412  
; CURRENT FILING DATE: 2003-04-30  
; NUMBER OF SEQ ID NOS: 429241  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15113  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-466-412-15113

Query Match 100.0%; Score 56; DB 32; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.46;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KALVOQMEQLRQ 12



Db 316 KALVQOMEQLRQ 327

RESULT 16  
US-09-079-030-123

; Sequence 123, Application US/09079030  
; GENERAL INFORMATION:  
; APPLICANT: Guevera, Jr., Juan G.  
; APPLICANT: Hoogeveen, Ron C.  
; APPLICANT: Moore, Paul J.  
; TITLE OF INVENTION: LIPOPROTEINS AS NUCLEIC ACID DELIVERY  
; TITLE OF INVENTION: VECTORS FOR TRANSFECTION OF EUKARYOTIC CELLS  
; NUMBER OF SEQUENCES: 229  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P. O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/079, 030  
; FILING DATE: Concurrently Herewith  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McMillian, Nabeela R.  
; REGISTRATION NUMBER: P-43,363  
; REFERENCE/DOCKET NUMBER: ARAG:003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 123:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 397 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; US-09-079-030-123

Query Match 100.0%; Score 56; DB 14; Length 397;  
Best Local Similarity 100.0%; Pred. No. 0.46;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVQOMEQLRQ 12  
Db 317 KALVQOMEQLRQ 328

RESULT 17  
US-09-791-84940  
; Sequence 84940, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biomomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Dauter, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791, 537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 84940  
; LENGTH: 401  
; TYPE: PRT  
; ORGANISM: Papio hamadryas anubis  
; US-09-791-84940

Query Match 100.0%; Score 56; DB 22; Length 401;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVQOMEQLRQ 12  
Db 300 KALVQOMEQLRQ 311

RESULT 18  
US-09-987-107-36  
; Sequence 36, Application US/09987107  
; GENERAL INFORMATION:  
; APPLICANT: GRAVERSEN, Jonas  
; APPLICANT: MOESTRUP, Soren  
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES  
; FILE REFERENCE: GRAVERSENIA  
; CURRENT APPLICATION NUMBER: US/09/987, 107  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/264, 022  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: DK PA2001 00057  
; PRIOR FILING DATE: 2001-01-15  
; PRIOR APPLICATION NUMBER: DK PA2000 01682  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 36  
; LENGTH: 401  
; TYPE: PRT  
; ORGANISM: Papio anubis  
; US-09-987-107-36

Query Match 100.0%; Score 56; DB 25; Length 401;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVQOMEQLRQ 12  
Db 300 KALVQOMEQLRQ 311

RESULT 19  
US-09-949-016-10854  
; Sequence 10854, Application US/09949016  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241, 755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237, 768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231, 498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10854  
; LENGTH: 426  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-949-016-10854

Query Match 100.0%; Score 56; DB 24; Length 426;  
Best Local Similarity 100.0%; Pred. No. 0.5;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVQOMEQLRQ 12

Db 346 KALVOQMEQLRQ 357

## RESULT 20

US-09-791-537-101059  
; Sequence 101059, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Dancer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 101059  
; LENGTH: 429  
; TYPE: PRT  
; ORGANISM: Macaca fascicularis  
US-09-791-537-101059

Query Match 100.0%; Score 56; DB 22; Length 429;  
Best Local Similarity 100.0%; Pred. No. 0.51;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12  
|||||  
Db 316 KALVOQMEQLRQ 327

## RESULT 21

US-09-987-107-34  
; Sequence 34, Application US/09987107  
; GENERAL INFORMATION:  
; APPLICANT: GRAVERSEN, Jonas  
; APPLICANT: MOESTRUP, Soren  
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES  
; FILE REFERENCE: GRAVERSEN1A  
; CURRENT APPLICATION NUMBER: US/09/987,107  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/264,022  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: DK PA2001 00057  
; PRIOR FILING DATE: 2001-01-15  
; PRIOR APPLICATION NUMBER: DK PA2000 01682  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 34  
; LENGTH: 429  
; TYPE: PRT  
; ORGANISM: Macaca fascicularis  
US-09-987-107-34

Query Match 100.0%; Score 56; DB 25; Length 429;  
Best Local Similarity 100.0%; Pred. No. 0.51;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12  
|||||  
Db 316 KALVOQMEQLRQ 327

## RESULT 22

PCT-US95-03660-10  
; Sequence 10, Application PC/TUS9503660  
; GENERAL INFORMATION:  
; APPLICANT: Tso, Patrick  
; TITLE OF INVENTION: EATING SUPPRESSANT PEPTIDES  
; NUMBER OF SEQUENCES: 87  
; CORRESPONDENCE ADDRESS:  
;

ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03660  
FILING DATE: 22-MAR-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 90212  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366  
TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US95-03660-10

Query Match 91.1%; Score 51; DB 1; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALVQQMEQLRQ 12  
|||||  
Db 1 ALVQQMEQLRQ 11

*Abandoned case*

RESULT 23  
US-08-216-537-10  
; Sequence 10, Application US/08216537  
; GENERAL INFORMATION:  
; APPLICANT: Tso, Patrick  
; TITLE OF INVENTION: EATING SUPPRESSANT PEPTIDES  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/216,537  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9021  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366  
TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 10:  
;

SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-216-537-10

Query Match 91.1%; Score 51; DB 6; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALVQOMEDLRQ 12  
DB 1 ALVQOMEDLRQ 11

*abandoned case*

RESULT 24  
US-08-408-858-10  
Sequence 10, Application US/08408858  
GENERAL INFORMATION:

APPLICANT: Tso, Patrick  
TITLE OF INVENTION: EATING SUPPRESSANT PEPTIDES  
NUMBER OF SEQUENCES: 87  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/408,858  
FILING DATE: 22-MAR-1995  
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: DIGILIO, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9021Z  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366

INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-408-858-10

Query Match 91.1%; Score 51; DB 8; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALVQOMEDLRQ 12  
DB 1 ALVQOMEDLRQ 11

*5840688*

RESULT 25  
US-09-791-139950  
Sequence 139950, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Biomomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 139950  
LENGTH: 376  
TYPE: PRT

US-09-791-537-139950

Query Match 91.1%; Score 51; DB 22; Length 376;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVQOMEDLR 11  
DB 296 KALVQOMEDLR 306

RESULT 26  
PCT-US01-10670-3  
Sequence 3, Application PC/TUS0110670  
GENERAL INFORMATION:

APPLICANT: Genaisance Pharmaceuticals, Inc.  
APPLICANT: Bentivegna, Steven C.  
APPLICANT: Choi, Julie Y.  
APPLICANT: Kilem, Stefanie E.  
APPLICANT: Koshy, Beena  
TITLE OF INVENTION: HAPLOTYPES OF THE APOA4 GENE  
FILE REFERENCE: MMH-0506PCT APOA4  
CURRENT APPLICATION NUMBER: PCT/US01/10670  
CURRENT FILING DATE: 2001-04-03  
PRIOR APPLICATION NUMBER: 60/194,362  
PRIOR FILING DATE: 2000-04-05  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 3  
LENGTH: 396  
TYPE: PRT

ORGANISM: Homo sapiens  
PCT-US01-10670-3

Query Match 91.1%; Score 51; DB 1; Length 396;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVQOMEDLR 11  
DB 346 KALVQOMEDLR 326

RESULT 27  
US-09-791-537-142101  
Sequence 142101, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Biomomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 142101  
LENGTH: 396  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-791-537-142101

Query Match 91.1%; Score 51; DB 22; Length 396;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLR 11  
|||  
Db 316 KALVOQMEQLR 326

RESULT 28  
US-09-802-640-16  
; Sequence 16, Application US/09802640  
; GENERAL INFORMATION:  
; APPLICANT: Braun, Andreas  
; APPLICANT: Bonsel Aruna  
; APPLICANT: Kley, Patrick

; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: 24736-2048  
; CURRENT APPLICATION NUMBER: US/09/802,640  
; CURRENT FILING DATE: 2001-03-09  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: FaestSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-802-640-16

Query Match 91.1%; Score 51; DB 23; Length 396;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLR 11  
|||  
Db 316 KALVOQMEQLR 326

RESULT 29  
US-10-403-902A-16  
; Sequence 16, Application US/10403902A  
; GENERAL INFORMATION:  
; APPLICANT: Braun, Andreas  
; APPLICANT: Bonsel, Aruna  
; APPLICANT: Kley, Patrick  
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: 24736-2048B  
; CURRENT APPLICATION NUMBER: US/10/403,902A  
; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: 09/802,640  
; PRIOR FILING DATE: 2001-03-09  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: FaestSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-403-902A-16

Query Match 91.1%; Score 51; DB 30; Length 396;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLR 11  
|||  
Db 316 KALVOQMEQLR 326

RESULT 30  
US-09-791-537-140603  
; Sequence 140603, Application US/09791537

; GENERAL INFORMATION:  
; APPLICANT: Biomix, Inc.  
; APPLICANT: Dede, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 140603  
; LENGTH: 391  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-140603

Query Match 87.5%; Score 49; DB 22; Length 391;  
Best Local Similarity 83.3%; Pred. No. 7.9;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLRQ 12  
|||  
Db 316 KALVOQMEQLRQ 327

RESULT 31  
US-09-800-729-208  
; Sequence 208, Application US/09800729  
; GENERAL INFORMATION:  
; APPLICANT: N1 et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 208  
; LENGTH: 391  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-800-729-208

Query Match 87.5%; Score 49; DB 23; Length 391;  
Best Local Similarity 83.3%; Pred. No. 7.9;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLRQ 12  
|||  
Db 316 KALVOQMEQLRQ 327

RESULT 32  
US-10-458-714-208  
; Sequence 208, Application US/10458714  
; GENERAL INFORMATION:  
; APPLICANT: N1 et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1  
; CURRENT APPLICATION NUMBER: US/10/458,714  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: US/09/800,729  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 208  
LENGTH: 391  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-458-714-208

Query Match 87.5%; Score 49; DB 30; Length 391;  
Best Local Similarity 83.3%; Pred. No. 7.9;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KALVOQMEOLRO 12  
Db 316 KALVOQLEOFRO 327

RESULT 33  
US-09-791-537-142580  
Sequence 142580, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Biomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 142580  
LENGTH: 395  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-791-537-142580

Query Match 87.5%; Score 49; DB 22; Length 395;  
Best Local Similarity 83.3%; Pred. No. 8;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KALVOQMEOLRO 12  
Db 316 KALVOQLEOFRO 327

RESULT 34  
US-09-987-107-35  
Sequence 35, Application US/09987107  
GENERAL INFORMATION:  
APPLICANT: GRAVERSEN, Jonas  
APPLICANT: MOESTRUP, Soren

TITLE OF INVENTION: APOLIPROTEINS ANALOGUES  
FILE REFERENCE: GRAVERSENIA  
CURRENT APPLICATION NUMBER: US/09/987,107  
CURRENT FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: US 60/264,022  
PRIOR FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: DK PA2001 00057  
PRIOR FILING DATE: 2001-01-15  
PRIOR APPLICATION NUMBER: DK PA2000 01682  
PRIOR FILING DATE: 2000-11-10  
NUMBER OF SEQ ID NOS: 91  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 35  
LENGTH: 395  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-987-107-35

Query Match 87.5%; Score 49; DB 25; Length 395;  
Best Local Similarity 83.3%; Pred. No. 8;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KALVOQMEOLRO 12

Db 316 KALVOQLEOFRO 327

RESULT 35  
US-09-791-537-139606  
Sequence 139606, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Biomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 139606  
LENGTH: 399  
TYPE: PRT  
ORGANISM: Mus musculus castaneus  
US-09-791-537-139606

Query Match 87.5%; Score 49; DB 22; Length 399;  
Best Local Similarity 83.3%; Pred. No. 8.1;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KALVOQMEOLRO 12  
Db 316 KALVOQLEOFRO 327

RESULT 36  
US-09-791-537-140880  
Sequence 140880, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Biomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 140880  
LENGTH: 435  
TYPE: PRT  
ORGANISM: Mus musculus castaneus  
US-09-791-537-140880

Query Match 87.5%; Score 49; DB 22; Length 435;  
Best Local Similarity 83.3%; Pred. No. 9;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KALVOQMEOLRO 12  
Db 316 KALVOQLEOFRO 327

RESULT 37  
US-09-791-537-142695  
Sequence 142695, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Biomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 142695  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Sus scrofa  
US-09-791-537-142695

Query Match 83.9%; Score 47; DB 22; Length 382;  
Best Local Similarity 83.3%; Pred. No. 17;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12  
|||:|  
Db 316 KALVOQVEDLRQ 327

RESULT 38  
US-09-800-729-206  
; Sequence 206, Application US/09800729  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 206  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-800-729-206

Query Match 83.9%; Score 47; DB 23; Length 382;  
Best Local Similarity 83.3%; Pred. No. 17;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12  
|||:|  
Db 316 KALVOQVEDLRQ 327

RESULT 39  
US-09-987-107-37  
; Sequence 37, Application US/09987107.  
; GENERAL INFORMATION:  
; APPLICANT: GRAVERSEN, Jonas  
; APPLICANT: MOESTRUP, Soren  
; TITLE OF INVENTION: APOLIPROTEINS ANALOGUES  
; FILE REFERENCE: GRAVERSENIA  
; CURRENT APPLICATION NUMBER: US/09/987,107  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/264,022  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: DK PA2001 00057  
; PRIOR FILING DATE: 2001-01-15  
; PRIOR APPLICATION NUMBER: DK PA2000 01682  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 37  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Sus scrofa  
US-09-987-107-37

Query Match 83.9%; Score 47; DB 25; Length 382;

Best Local Similarity 83.3%; Pred. No. 17;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12  
|||:|  
Db 316 KALVOQVEDLRQ 327

RESULT 40  
US-10-458-714-206  
; Sequence 206, Application US/10458714  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1  
; CURRENT APPLICATION NUMBER: US/10/458,714  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: US/09/800,729  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 206  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-458-714-206

Query Match 83.9%; Score 47; DB 30; Length 382;  
Best Local Similarity 83.3%; Pred. No. 17;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12  
|||:|  
Db 316 KALVOQVEDLRQ 327

Search completed: September 8, 2003, 16:10:24  
Job time : 172 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2003, 16:05:28 ; Search time 9 seconds  
(without alignments)  
12.377 Million cell updates/sec

Title: US-09-991-809-1

Perfect score: 56

Sequence: 1 KALVQOMEQLRQ 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 44440 seqs, 9282514 residues

Total number of hits satisfying chosen parameters: 44440

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	100.0	396	5	US-09-842-364A-8
2	47	83.9	382	5	US-09-842-364A-7
3	35	62.5	293	5	US-09-897-516A-8027
4	35	62.5	4827	7	US-60-487-610-2167
5	33	58.9	941	5	US-09-581-286A-406
6	33	58.9	941	7	US-60-495-589-319
7	33	58.9	945	5	US-09-581-286A-282
8	32	57.1	126	1	PCT-US02-37235-98
9	32	57.1	145	5	US-09-674-546A-2589
10	32	57.1	373	5	US-09-674-546A-1871
11	32	57.1	458	7	US-60-495-114-1391
12	32	57.1	558	5	US-09-674-546A-1873
13	32	57.1	558	5	US-09-674-546A-1875
14	32	57.1	1329	7	US-60-490-890-1670
15	32	57.1	1402	7	US-60-490-890-1831
16	32	57.1	4574	7	US-60-479-073-497
17	32	57.1	4684	1	PCT-US03-14382-432
18	32	57.1	4684	1	PCT-US03-14382-434
19	32	57.1	4684	1	PCT-US03-14382A-432
20	32	57.1	4684	7	US-60-479-073-497
21	32	57.1	4684	7	US-60-479-073-506
22	32	57.1	4684	7	US-60-479-073-506
23	31	55.4	180	5	US-09-897-516A-7884
24	31	55.4	150	5	US-09-897-516A-6463
25	31	55.4	384	1	PCT-US03-20001-42
26	31	55.4	585	7	US-60-490-890-1243

27	31	55.4	589	5	US-09-897-516A-6818	Sequence 6818, Ap
28	31	55.4	702	1	PCT-US03-24669-288	Sequence 288, App
29	31	55.4	802	1	PCT-US03-23245-30	Sequence 30, App1
30	31	55.4	1069	1	PCT-US03-23245-22	Sequence 22, App1
31	31	55.4	1164	7	US-60-487-610-2411	Sequence 2411, Ap
32	31	55.4	1164	7	US-60-485-450-1525	Sequence 1525, Ap
33	31	55.4	1284	7	US-60-487-610-1521	Sequence 1521, Ap
34	31	55.4	1284	7	US-60-485-450-1521	Sequence 1521, Ap
35	31	55.4	1427	7	US-60-490-890-824	Sequence 824, App
36	31	55.4	1464	1	PCT-US02-18638A-36	Sequence 36, App1
37	31	55.4	1464	7	US-60-487-610-2409	Sequence 2409, Ap
38	31	55.4	1464	7	US-60-485-450-1523	Sequence 1523, Ap
39	31	55.4	2101	5	US-09-296-662B-32	Sequence 32, App1
40	31	55.4	2115	5	US-09-296-662B-33	Sequence 33, App1
41	31	55.4	3006	1	PCT-US02-24483-36	Sequence 36, App1
42	31	55.4	5183	7	US-60-490-890-719	Sequence 719, App
43	30.5	54.5	2548	5	US-09-851-682B-1	Sequence 1, App1
44	30	53.6	218	5	US-09-897-516A-6025	Sequence 6025, Ap
45	30	53.6	381	7	US-60-487-610-2289	Sequence 2289, Ap

## ALIGNMENTS

```
RESULT 1
US-09-842-364A-8
; Sequence 8, Application US/09842364A
; GENERAL INFORMATION:
; APPLICANT: Yen-Potin, Frances
; APPLICANT: Denison, Blake
; APPLICANT: Bour, Barbara
; APPLICANT: Bihain, Bernard
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Ducleter, Aymeric
; TITLE OF INVENTION: Apolipoprotein A-IV-related Protein: Polypeptide, Polynucleotide
; FILE REFERENCE: G-089US04CIP
; CURRENT APPLICATION NUMBER: US/09/842,364A
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 09/599,362
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/141,032
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: PCT/IB99/02058
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: US 09/469,099
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: US 60/113,686
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Human Apoa IV
US-09-842-364A-8

Query Match      100.0%; Score 56; DB 5; Length 396;
Best Local Similarity 100.0%; Pred. No. 0.0045;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 KALVQOMEQLRQ 12
Db      316 KALVQOMEQLRQ 327

RESULT 2
US-09-842-364A-7
; Sequence 7, Application US/09842364A
; GENERAL INFORMATION:
; APPLICANT: Yen-Potin, Frances
; APPLICANT: Denison, Blake
```

```

; APPLICANT: Bour, Barbara
; APPLICANT: Bihain, Bernard
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclet, Aymeric
; APPLICANT: Bouguet, Lydie
; TITLE OF INVENTION: Apolipoprotein A-IV-related Protein: Polypeptide, Polynucleotide
; FILE REFERENCE: G-089US04CIP
; CURRENT APPLICATION NUMBER: US/09/842,364A
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 09/599,362
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/141,032
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: PCT/IB99/02058
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: US 09/469,099
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: US 60/113,686
; PRIOR FILING DATE: 1998-12-22
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 7
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Swine Apca IV
US-09-842-364A-7
```

Query Match 83.9%; Score 47; DB 5; Length 382;

Best Local Similarity 83.3%; Pred. No. 0.16; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 KALVQOQEDLRQ 12
    |||||:||||
Db 316 KALVQOVEDLRQ 327
```

RESULT 3

US-09-897-516A-8027

Sequence 8027, Application US/09897516A

GENERAL INFORMATION:

```

; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huebner, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Kraasoml-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spilidonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 8027
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516A-8027
```

Query Match 62.5%; Score 35; DB 5; Length 293;

Best Local Similarity 58.3%; Pred. No. 15; Mismatches 3; Indels 2; Gaps 0;

```

QY 1 KALVQOQEDLRQ 12
    |||||:||||
Db 67 KALVQOQEDLRQ 78
```

RESULT 4

US-60-487-610-2167

```

; Sequence 2167, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2167
; LENGTH: 4827
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-2167
```

Query Match 62.5%; Score 35; DB 7; Length 4827;

Best Local Similarity 77.8%; Pred. No. 3.6e+02; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 KALVQOQEO 9
    |||||:||||
Db 2222 KALVQOQEO 2230
```

RESULT 5

US-09-581-286A-406

Sequence 406, Application US/09581286A

GENERAL INFORMATION:

```

; APPLICANT: ROSS, BRUCE C.
; APPLICANT: BARR, IAN G.
; APPLICANT: PATTERSON, MICHELLE A.
; APPLICANT: AGIUS, CATHERINE T.
; APPLICANT: ROTHEL, LINDA J.
; APPLICANT: MARGITTS, MAL B.
; APPLICANT: HOCKING, DIANNA M.
; APPLICANT: WEBB, ELIZABETH A.
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS POLYPEPTIDES AND NUCLEOTIDES
; FILE REFERENCE: 4137-3
; CURRENT APPLICATION NUMBER: US/09/581,286A
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: PCT/AU98/01023
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: AU PP 0839
; PRIOR FILING DATE: 1997-12-10
; PRIOR APPLICATION NUMBER: AU PP 1182
; PRIOR FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: AU PP 1846
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: AU PP 2264
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: AU PP 2911
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: AU PP 3128
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: AU PP 3338
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: AU PP 3654
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: AU PP 4917
; PRIOR FILING DATE: 1998-07-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 406
; LENGTH: 941
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-581-286A-406
```

Query Match 58.9%; Score 33; DB 5; Length 941;



Best Local Similarity 63.6%; Pred. No. 1.2e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ALVQOMEQLRQ 12  
|||::|  
Db 370 ALVAIEISLRQ 380

RESULT 6  
US-60-495-589-319

; Sequence 319, Application US/60495589 -  
; GENERAL INFORMATION:  
; APPLICANT: Proguiske-Fox, Ann  
; APPLICANT: Hillman, Jeffrey D.  
; APPLICANT: Handfield, Martin  
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTIDE  
; FILE REFERENCE: 02-042  
; CURRENT APPLICATION NUMBER: US/60/495,589  
; CURRENT FILING DATE: 2003-08-15  
; NUMBER OF SEQ ID NOS: 354  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 319  
; LENGTH: 941  
; TYPE: PRT  
; ORGANISM: Porphyromonas gingivalis  
US-60-495-589-319

Query Match 58.9%; Score 33; DB 7; Length 941;  
Best Local Similarity 63.6%; Pred. No. 1.2e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ALVQOMEQLRQ 12  
|||::|  
Db 370 ALVAIEISLRQ 380

RESULT 7  
US-09-581-286A-282

; Sequence 282, Application US/09581286A  
; GENERAL INFORMATION:  
; APPLICANT: ROSS, BRUCE C.  
; APPLICANT: BARR, IAN G.  
; APPLICANT: PATTERSON, MICHELLE A.  
; APPLICANT: AGIUS, CATHERINE T.  
; APPLICANT: ROTHEN, LINDA J.  
; APPLICANT: MARGETS, MAL B.  
; APPLICANT: HOCKING, DIANNA M.  
; APPLICANT: WEBB, ELIZABETH A.  
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS POLYPEPTIDES AND NUCLEOTIDES  
; FILE REFERENCE: 4137-3  
; CURRENT APPLICATION NUMBER: US/09/581,286A  
; CURRENT FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: PCT/AU98/01023  
; PRIOR FILING DATE: 1998-12-10  
; PRIOR APPLICATION NUMBER: AU PP 0839  
; PRIOR FILING DATE: 1997-12-10  
; PRIOR APPLICATION NUMBER: AU PP 1182  
; PRIOR FILING DATE: 1997-12-31  
; PRIOR APPLICATION NUMBER: AU PP 1846  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: AU PP 2264  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: AU PP 2911  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: AU PP 3128  
; PRIOR FILING DATE: 1998-04-23  
; PRIOR APPLICATION NUMBER: AU PP 3338  
; PRIOR FILING DATE: 1998-05-05  
; PRIOR APPLICATION NUMBER: AU PP 3654  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: AU PP 4917  
; PRIOR FILING DATE: 1998-07-29

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 721  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 282  
; LENGTH: 945  
; TYPE: PRT  
; ORGANISM: Porphyromonas gingivalis  
US-09-581-286A-282

Query Match 58.9%; Score 33; DB 5; Length 945;  
Best Local Similarity 63.6%; Pred. No. 1.2e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ALVQOMEQLRQ 12  
|||::|  
Db 374 ALVAIEISLRQ 384

RESULT 8  
PCT-US02-37235-98

; Sequence 98, Application PC/TUS0237235  
; GENERAL INFORMATION:  
; APPLICANT: Handfield, Martin  
; APPLICANT: Hillman, Jeffrey  
; APPLICANT: Proguiske-Fox, Ann  
; TITLE OF INVENTION: Identification of Actinobacillus actinomycetemcomitans Antigens f  
; FILE REFERENCE: MHH01-662A  
; CURRENT APPLICATION NUMBER: PCT/US02/37235  
; CURRENT FILING DATE: 2003-06-16  
; NUMBER OF SEQ ID NOS: 234  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 98  
; LENGTH: 126  
; TYPE: PRT  
; ORGANISM: Actinobacillus actinomycetemcomitans  
PCT-US02-37235-98

Query Match 57.1%; Score 32; DB 1; Length 126;  
Best Local Similarity 70.0%; Pred. No. 19;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LVQOMEQLRQ 12  
|||::|  
Db 85 LVKOMEQASQ 94

RESULT 9  
US-09-674-546A-2589

; Sequence 2589, Application US/09674546A  
; GENERAL INFORMATION:  
; APPLICANT: Institute for Genomic Research  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Neisseria Meningitidis Antigens and Compositions  
; FILE REFERENCE: CHIR-0334  
; CURRENT APPLICATION NUMBER: US/09/674,546A  
; CURRENT FILING DATE: 1999-04-30  
; NUMBER OF SEQ ID NOS: 3264  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 2589  
; LENGTH: 145  
; TYPE: PRT  
; ORGANISM: Neisseria meningitidis  
US-09-674-546A-2589

Query Match 57.1%; Score 32; DB 5; Length 145;  
Best Local Similarity 54.5%; Pred. No. 22;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KALVQOMEQLR 11  
|||::|  
Db 103 QAPLQAVEQLR 113

RESULT 10  
US-09-674-546A-1871  
; Sequence 1871, Application US/09674546A  
; GENERAL INFORMATION:  
; APPLICANT: Institute for Genomic Research  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Neisseria Meningitidis Antigens and Compositions  
; FILE REFERENCE: CHIR-0334  
; CURRENT APPLICATION NUMBER: US/09/674,546A  
; CURRENT FILING DATE: 1999-04-30  
; NUMBER OF SEQ ID NOS: 3264  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO: 1871  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-09-674-546A-1871

Query Match 57.1%; Score 32; DB 5; Length 373;  
Best Local Similarity 50.0%; Pred. No. 64;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 KALVQOMEQLRQ 12  
||:|:|:|:|  
Db 85 KAMKQELWVRQ 96

NS

RESULT 11  
US-60-495-114-1391  
; Sequence 1391, Application US/60495114  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND  
; FILE REFERENCE: CL001480  
; CURRENT APPLICATION NUMBER: US/60/495,114  
; CURRENT FILING DATE: 2003-08-15  
; NUMBER OF SEQ ID NOS: 91238  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 1391  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-495-114-1391

Query Match 57.1%; Score 32; DB 7; Length 458;  
Best Local Similarity 50.0%; Pred. No. 81;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 KALVQOMEQLRQ 12  
||:|:|:|:|  
Db 329 EALIQFLQVHQ 340

RESULT 12  
US-09-674-546A-1873  
; Sequence 1873, Application US/09674546A  
; GENERAL INFORMATION:  
; APPLICANT: Institute for Genomic Research  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Neisseria Meningitidis Antigens and Compositions  
; FILE REFERENCE: CHIR-0334  
; CURRENT APPLICATION NUMBER: US/09/674,546A  
; CURRENT FILING DATE: 1999-04-30  
; NUMBER OF SEQ ID NOS: 3264  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO: 1873  
; LENGTH: 558  
; TYPE: PRT  
; ORGANISM: Neisseria meningitidis  
US-09-674-546A-1873

Query Match 57.1%; Score 32; DB 5; Length 558;  
Best Local Similarity 50.0%; Pred. No. 1e+02;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 KALVQOMEQLRQ 12  
||:|:|:|:|  
Db 270 KAMKQELWVRQ 281

NS

RESULT 13  
US-09-674-546A-1875  
; Sequence 1875, Application US/09674546A  
; GENERAL INFORMATION:  
; APPLICANT: Institute for Genomic Research  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Neisseria Meningitidis Antigens and Compositions  
; FILE REFERENCE: CHIR-0334  
; CURRENT APPLICATION NUMBER: US/09/674,546A  
; CURRENT FILING DATE: 1999-04-30  
; NUMBER OF SEQ ID NOS: 3264  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO: 1875  
; LENGTH: 558  
; TYPE: PRT  
; ORGANISM: Neisseria meningitidis  
US-09-674-546A-1875

Query Match 57.1%; Score 32; DB 5; Length 558;  
Best Local Similarity 50.0%; Pred. No. 1e+02;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 KALVQOMEQLRQ 12  
||:|:|:|:|  
Db 270 KAMKQELWVRQ 281

NS

RESULT 14  
US-60-490-890-1670  
; Sequence 1670, Application US/60490890  
; GENERAL INFORMATION:  
; APPLICANT: Li, Martha  
; APPLICANT: Rupnow, Brent A.  
; APPLICANT: Webster, Kevin R.  
; APPLICANT: Jackson, Donald  
; APPLICANT: Wong, Tai W.  
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION  
; FILE REFERENCE: D0310 P5P  
; CURRENT APPLICATION NUMBER: US/60/490,890  
; CURRENT FILING DATE: 2003-07-29  
; NUMBER OF SEQ ID NOS: 2779  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO: 1670  
; LENGTH: 1329  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-490-890-1670

Query Match 57.1%; Score 32; DB 7; Length 1329;  
Best Local Similarity 58.3%; Pred. No. 2.7e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KALVQOMEQLRQ 12  
||:|:|:|:|  
Db 1003 KALQOQOQOQOQO 1014

RESULT 15  
US-60-490-890-1831  
; Sequence 1831, Application US/60490890  
; GENERAL INFORMATION:  
; APPLICANT: Li, Martha  
; APPLICANT: Rupnow, Brent A.

```
; APPLICANT: Webster, Kevin R.
; APPLICANT: Jackson, Donald
; APPLICANT: Wong, Tai W.
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 PSP
; CURRENT APPLICATION NUMBER: US/60/490,890
; CURRENT FILING DATE: 2003-07-29
; NUMBER OF SEQ ID NOS: 2779
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1831
; LENGTH: 1402
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-60-490-890-1831
```

```
Query Match          57.1%; Score 32; DB 7; Length 1402;
Best Local Similarity 63.6%; Pred. No. 2.9e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      2 ALVQOQEQLRQ 12
      |||||
      419 ALASQLEQERQ 429
```

```
RESULT 16
; US-60-479-073-495
; Sequence 495, Application US/60479073
; GENERAL INFORMATION:
```

```
; APPLICANT: De Wilde, Gert Jules Hector
; APPLICANT: Saunders, Michael John Scott
; APPLICANT: Logghe, Marc Georges
; TITLE OF INVENTION: Amino acid sequences useful for developing compounds for the
; TITLE OF INVENTION: prevention and/or treatment of metabolic diseases and nucleotide
; FILE REFERENCE: D00590,70042.US
; CURRENT APPLICATION NUMBER: US/60/479,073
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 526
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 495
; LENGTH: 4574
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-60-479-073-495
```

```
Query Match          57.1%; Score 32; DB 7; Length 4574;
Best Local Similarity 87.5%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      5 QOQEQLRQ 12
      |||||
      2566 QOQEQERQ 2573
```

```
RESULT 17
; PCT-US03-14382-432
; Sequence 432, Application PC/TUS0314382
; GENERAL INFORMATION:
```

```
; APPLICANT: Emory University
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; FILE REFERENCE: 12804-011W01
; CURRENT APPLICATION NUMBER: PCT/US03/14382
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: US 60/378,711
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 432
; LENGTH: 4684
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US03-14382-432
```

```
Query Match          57.1%; Score 32; DB 1; Length 4684;
Best Local Similarity 87.5%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      5 QOQEQLRQ 12
      |||||
      2676 QOQEQERQ 2683
```

```
RESULT 18
; PCT-US03-14382-434
```

```
; Sequence 434, Application PC/TUS0314382
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; FILE REFERENCE: 12804-011W01
; CURRENT APPLICATION NUMBER: PCT/US03/14382
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: US 60/378,711
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 434
; LENGTH: 4684
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US03-14382-434
```

```
Query Match          57.1%; Score 32; DB 1; Length 4684;
Best Local Similarity 87.5%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      5 QOQEQLRQ 12
      |||||
      2676 QOQEQERQ 2683
```

```
RESULT 19
; PCT-US03-14382A-432
; Sequence 432, Application PC/TUS0314382A
; GENERAL INFORMATION:
```

```
; APPLICANT: Emory University
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; FILE REFERENCE: 12804-011W01
; CURRENT APPLICATION NUMBER: PCT/US03/14382A
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US 60/378,711
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 432
; LENGTH: 4684
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US03-14382A-432
```

```
Query Match          57.1%; Score 32; DB 1; Length 4684;
Best Local Similarity 87.5%; Pred. No. 1.1e+03;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      5 QOQEQLRQ 12
      |||||
      2676 QOQEQERQ 2683
```

```
RESULT 20
; PCT-US03-14382A-434
; Sequence 434, Application PC/TUS0314382A
; GENERAL INFORMATION:
; APPLICANT: Emory University
; PCT-US03-14382-432
```

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: ANTIVIRAL AGENTS
; FILE REFERENCE: 12804-011W01
; CURRENT APPLICATION NUMBER: PCT/US03/14382A
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US 60/378,711
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 434
; LENGTH: 4684
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-14382A-434

Query Match
Best Local Similarity 57.1%; Score 32; DB 1; Length 4684;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 QOMEQLRQ 12
Db 2676 QOMEQLRQ 2683

RESULT 21
US-60-479-073-497
; Sequence 497, Application US/60479073
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Gert Jules Hector
; APPLICANT: Sanders, Michael John Scott
; TITLE OF INVENTION: Amino acid sequences useful for developing compounds for the
; TITLE OF INVENTION: prevention and/or treatment of metabolic diseases and nucleotide
; FILE REFERENCE: D00590.70042.US
; CURRENT APPLICATION NUMBER: US/60/479,073
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 526
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 497
; LENGTH: 4684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-479-073-497

Query Match
Best Local Similarity 57.1%; Score 32; DB 7; Length 4684;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 QOMEQLRQ 12
Db 2676 QOMEQLRQ 2683

RESULT 22
US-60-479-073-506
; Sequence 506, Application US/60479073
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Gert Jules Hector
; APPLICANT: Sanders, Michael John Scott
; TITLE OF INVENTION: Amino acid sequences useful for developing compounds for the
; TITLE OF INVENTION: prevention and/or treatment of metabolic diseases and nucleotide
; FILE REFERENCE: D00590.70042.US
; CURRENT APPLICATION NUMBER: US/60/479,073
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 526
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 506
; LENGTH: 4684
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```

US-60-479-073-506

Query Match
Best Local Similarity 57.1%; Score 32; DB 7; Length 4684;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 QOMEQLRQ 12
Db 2676 QOMEQLRQ 2683

RESULT 23
US-09-897-516A-7884
; Sequence 7884, Application US/09897516A
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Hueising, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Kraasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 2001-06-29
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 7884
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516A-7884

Query Match
Best Local Similarity 55.4%; Score 31; DB 5; Length 80;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALVQOMEQLRQ 12
Db 50 SLVRENEQLKQ 60

RESULT 24
US-09-897-516A-6463
; Sequence 6463, Application US/09897516A
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Hueising, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Kraasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 6463
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516A-6463

Query Match
Best Local Similarity 55.4%; Score 31; DB 5; Length 150;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

OY 1 KALVOQMEQLRQ 12  
||| :| :| :|  
DB 105 KALAGEMGTIRQ 116

## RESULT 25

PCT-US03-20001-42  
; Sequence 42, Application PC/TUS0320001  
; GENERAL INFORMATION:  
; APPLICANT: Boyes, Douglas  
; APPLICANT: Davis, Keith  
; APPLICANT: Jones, Alan  
; APPLICANT: Ullah, Hemayet  
; APPLICANT: Chen, Jin-Gui  
; APPLICANT: Mulputi, Rao  
; APPLICANT: Chatterjee, Ani  
; APPLICANT: Waid, Mary  
; TITLE OF INVENTION: METHODS FOR IMPROVING PLANT AGRONOMICAL TRAITS BY ALTERING THE  
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY OF PLANT G-PROTEIN ALPHA AND BETA SUBUNIT  
; FILE REFERENCE: 2155US  
; CURRENT APPLICATION NUMBER: PCT/US03/20001  
; CURRENT FILING DATE: 2003-06-24  
; PRIOR APPLICATION NUMBER: 60/392,730  
; PRIOR FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: 60/445,208  
; PRIOR FILING DATE: 2003-02-05  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 42  
; LENGTH: 384  
; TYPE: PRT  
; ORGANISM: Nicotiana tabacum  
PCT-US03-20001-42

Query Match 55.4%; Score 31; DB 1; Length 384;  
Best Local Similarity 58.3%; Pred. No. 99;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 KALVOQMEQLRQ 12  
||| :| :| :|  
DB 134 KDLVODIEALRK 145

## RESULT 26

US-60-490-890-1243  
; Sequence 1243, Application US/60490890  
; GENERAL INFORMATION:  
; APPLICANT: Li, Martha  
; APPLICANT: Rupnow, Brent A.  
; APPLICANT: Webster, Kevin R.  
; APPLICANT: Jackson, Donald  
; APPLICANT: Wong, Tai W.  
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION  
; FILE REFERENCE: D0310 PSP  
; CURRENT APPLICATION NUMBER: US/60/490,890  
; CURRENT FILING DATE: 2003-07-29  
; NUMBER OF SEQ ID NOS: 2779  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1243  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-490-890-1243

Query Match 55.4%; Score 31; DB 7; Length 585;  
Best Local Similarity 66.7%; Pred. No. 1.6e+02;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 KALVOQMEQLRQ 12  
||| :| :| :|  
DB 404 KHLVQOEQLRQ 415

## RESULT 27

US-09-897-516A-6818  
; Sequence 6818, Application US/09897516A  
; GENERAL INFORMATION:  
; APPLICANT: Corbin, David R.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Hueising, Joseph E.  
; APPLICANT: Malvar, Thomas M.  
; APPLICANT: Krasomil-Osterfeld, Karina C.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Spiridonov, Sergei  
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
; FILE REFERENCE: 38-21(51847)B  
; CURRENT APPLICATION NUMBER: US/09/897,516A  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/215,161  
; PRIOR FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 8415  
; SEQ ID NO 6818  
; LENGTH: 589  
; TYPE: PRT  
; ORGANISM: Xenorhabdus sp.  
US-09-897-516A-6818

Query Match 55.4%; Score 31; DB 5; Length 589;  
Best Local Similarity 58.3%; Pred. No. 1.6e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 KALVOQMEQLRQ 12  
||| :| :| :|  
DB 512 KALSRATNQLRQ 523

## RESULT 28

PCT-US03-24669-288  
; Sequence 288, Application PC/TUS0324669  
; GENERAL INFORMATION:  
; APPLICANT: diadexus, Inc.  
; APPLICANT: Macina, Roberto  
; APPLICANT: Salceda, Susana  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Sun, Yongming  
; APPLICANT: Turner, Leah  
; TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and P  
; FILE REFERENCE: DEX-0443  
; CURRENT APPLICATION NUMBER: PCT/US03/24669  
; CURRENT FILING DATE: 2003-08-06  
; PRIOR APPLICATION NUMBER: US 60/401,469  
; PRIOR FILING DATE: 2002-08-06  
; NUMBER OF SEQ ID NOS: 396  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 288  
; LENGTH: 702  
; TYPE: PRT  
; ORGANISM: Homo sapien  
PCT-US03-24669-288

Query Match 55.4%; Score 31; DB 1; Length 702;  
Best Local Similarity 54.5%; Pred. No. 2e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 KALVOQMEQLR 11  
||| :| :| :|  
DB 585 KSLSQOIEINR 595

## RESULT 29

PCT-US03-23245-30  
; Sequence 30, Application PC/TUS0323245  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE CORPORATION; MASON, Patricia M.;

```

; APPLICANT: SWARNAKAR, Anita; JIANG, Xin;
; APPLICANT: JACKSON, Alan A.; KABLE, Amy E.;
; APPLICANT: TANG, Y. Tom; FORSYTHE, Ian J.;
; APPLICANT: ELIOTT, Vicki S.; LEE, Soo Yeun;
; APPLICANT: LEE, Ernestine A.; ISON, Craig H.;
; APPLICANT: HAFALIA, April J.A.; KHARE, Reena;
; APPLICANT: MARQUIS, Joseph P.; BECHA, Shanya D.;
; APPLICANT: BULLOCH, Sean A.; BLAKE, Julie J.;
; APPLICANT: GANDHI, Ameena; GRIFPIN, Jennifer A.;
; APPLICANT: LEE, Sally; YUE, Henry;
; APPLICANT: YANG, Yonghong G.; SPRAGUE, William W.;
; APPLICANT: BAUGHN, Mariah R.; WANG, Jonathan T.;
; APPLICANT: GERA, Mili; GIETZEN, Kimberly J.;
; APPLICANT: NGUYEN, Daniel B.; LU, Dyang Aina M.
; TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-1472 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/23245
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: US 60/398,907
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/407,068
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US 60/414,139
; PRIOR FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/424,094
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/440,912
; PRIOR FILING DATE: 2003-01-17
; PRIOR APPLICATION NUMBER: US 60/442,419
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PERL Program
; SEQ ID NO 30
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 7523973CD1
PCT-US03-23245-30

Query Match          55.4%; Score 31; DB 1; Length 802;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 LVQOMEQLRQ 12
DB 635 LIQQLISQLAQ 644
```

```

; PRIOR APPLICATION NUMBER: US 60/398,907
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/407,068
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US 60/414,139
; PRIOR FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/424,094
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/440,912
; PRIOR FILING DATE: 2003-01-17
; PRIOR APPLICATION NUMBER: US 60/442,419
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PERL Program
; SEQ ID NO 22
; LENGTH: 1069
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 7519735CD1
PCT-US03-23245-22

Query Match          55.4%; Score 31; DB 1; Length 1069;
Best Local Similarity 63.6%; Pred. No. 3.2e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALVQOMEQLRQ 12
DB 496 ALQQLFHLQQLQ 506
```

```

RESULT 30
PCT-US03-23245-22
; Sequence 22, Application PC/TUS0323245
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; MASON, Patricia M.;
; APPLICANT: SWARNAKAR, Anita; JIANG, Xin;
; APPLICANT: JACKSON, Alan A.; KABLE, Amy E.;
; APPLICANT: TANG, Y. Tom; FORSYTHE, Ian J.;
; APPLICANT: ELIOTT, Vicki S.; LEE, Soo Yeun;
; APPLICANT: LEE, Ernestine A.; ISON, Craig H.;
; APPLICANT: HAFALIA, April J.A.; KHARE, Reena;
; APPLICANT: MARQUIS, Joseph P.; BECHA, Shanya D.;
; APPLICANT: BULLOCH, Sean A.; BLAKE, Julie J.;
; APPLICANT: GANDHI, Ameena; GRIFPIN, Jennifer A.;
; APPLICANT: LEE, Sally; YUE, Henry;
; APPLICANT: YANG, Yonghong G.; SPRAGUE, William W.;
; APPLICANT: BAUGHN, Mariah R.; WANG, Jonathan T.;
; APPLICANT: GERA, Mili; GIETZEN, Kimberly J.;
; APPLICANT: NGUYEN, Daniel B.; LU, Dyang Aina M.
; TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-1472 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/23245
; CURRENT FILING DATE: 2003-07-23
```

```

RESULT 31
US-60-487-610-2411
; Sequence 2411, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2411
; LENGTH: 1164
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-2411

Query Match          55.4%; Score 31; DB 7; Length 1164;
Best Local Similarity 54.5%; Pred. No. 3.5e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KALVQOMEQLR 11
DB 936 KSLSQQLIENIR 946
```

```

RESULT 32
US-60-485-450-1525
; Sequence 1525, Application US/60485450
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01470
; CURRENT APPLICATION NUMBER: US/60/485,450
```

;; CURRENT FILING DATE: 2003-07-09  
;; NUMBER OF SEQ ID NOS: 47859  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 1525  
;; LENGTH: 1164  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-60-485-450-1525

Query Match 55.4%; Score 31; DB 7; Length 1164;  
Best Local Similarity 54.5%; Pred. No. 3.5e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KALVOMEQRLR 11  
|:|:|:|:|:  
Db 936 KSLSQOIEIR 946

ND

RESULT 33  
US-60-487-610-2407  
; Sequence 2407, Application US/60487610  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,  
; FILE REFERENCE: CLO01469  
; CURRENT APPLICATION NUMBER: US/60/487,610  
; CURRENT FILING DATE: 2003-07-17  
; NUMBER OF SEQ ID NOS: 97101  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2407  
; LENGTH: 1284  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-487-610-2407

Query Match 55.4%; Score 31; DB 7; Length 1284;  
Best Local Similarity 54.5%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KALVOMEQRLR 11  
|:|:|:|:|:  
Db 1056 KSLSQOIEIR 1066

ND

RESULT 34  
US-60-485-450-1521  
; Sequence 1521, Application US/60485450  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C  
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES  
; FILE REFERENCE: CLO01470  
; CURRENT APPLICATION NUMBER: US/60/485,450  
; CURRENT FILING DATE: 2003-07-09  
; NUMBER OF SEQ ID NOS: 47859  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1521  
; LENGTH: 1284  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-485-450-1521

Query Match 55.4%; Score 31; DB 7; Length 1284;  
Best Local Similarity 54.5%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KALVOMEQRLR 11

Db 1056 KSLSQOIEIR 1066

ND

RESULT 35  
US-60-490-890-824  
; Sequence 824, Application US/60490890  
; GENERAL INFORMATION:  
; APPLICANT: Li, Martha  
; APPLICANT: Rudnow, Brent A.  
; APPLICANT: Webster, Kevin R.  
; APPLICANT: Jackson, Donald  
; APPLICANT: Wong, Tai W.  
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION  
; FILE REFERENCE: D0310 PSP  
; CURRENT APPLICATION NUMBER: US/60/490,890  
; CURRENT FILING DATE: 2003-07-29  
; NUMBER OF SEQ ID NOS: 2779  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 824  
; LENGTH: 1427  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-490-890-824

Query Match 55.4%; Score 31; DB 7; Length 1427;  
Best Local Similarity 50.0%; Pred. No. 4.4e+02;  
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 3 LVQMEQRLR 12  
|:|:|:|:|:  
Db 1057 LQLELELRK 1066

ND

RESULT 36  
PCT-US02-18638A-36  
; Sequence 36, Application PC/TUS0218638A  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY  
; FILE REFERENCE: MRI-035PC  
; CURRENT APPLICATION NUMBER: PCT/US02/18638A  
; CURRENT FILING DATE: 2002-06-12  
; PRIOR APPLICATION NUMBER: US 60/298,159  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,155  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/335,936  
; PRIOR FILING DATE: 2001-11-14  
; NUMBER OF SEQ ID NOS: 238  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36  
; LENGTH: 1464  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-18638A-36

Query Match 55.4%; Score 31; DB 1; Length 1464;  
Best Local Similarity 54.5%; Pred. No. 4.5e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KALVOMEQRLR 11  
|:|:|:|:|:  
Db 1236 KSLSQOIEIR 1246

ND

RESULT 37  
US-60-487-610-2409  
; Sequence 2409, Application US/60487610  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele

APPLICANT: HUANG, Hongjin  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,  
TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: C0001469  
CURRENT APPLICATION NUMBER: US/60/487,610  
NUMBER OF SEQ ID NOS: 97101  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2409  
LENGTH: 1464  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-60-487-610-2409

Query Match 55.4%; Score 31; DB 7; Length 1464;  
Best Local Similarity 54.5%; Pred. No. 4.5e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLR 11  
Db 1236 KLSQIQENIR 1246

ND

RESULT 38  
US-60-485-450-1523  
Sequence 1523, Application US/60485450  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C  
TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES  
FILE REFERENCE: C0001470  
CURRENT APPLICATION NUMBER: US/60/485,450  
CURRENT FILING DATE: 2003-07-09  
NUMBER OF SEQ ID NOS: 47859  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1523  
LENGTH: 1464  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-60-485-450-1523

Query Match 55.4%; Score 31; DB 7; Length 1464;  
Best Local Similarity 54.5%; Pred. No. 4.5e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLR 11  
Db 1236 KLSQIQENIR 1246

ND

RESULT 39  
US-09-296-662B-32  
Sequence 32, Application US/09296662B  
GENERAL INFORMATION:  
APPLICANT: Anthony Rosen  
APPLICANT: Donald W. Nicholson  
APPLICANT: Livia Casciola-Rosen  
APPLICANT: Felipe A. Andrade  
APPLICANT: Sophie Roy  
APPLICANT: Nancy A. Thornberry  
TITLE OF INVENTION: AUTOANTIGENIC FRAGMENTS, METHODS AND  
TITLE OF INVENTION: ASSAYS  
FILE REFERENCE: 20221Y  
CURRENT APPLICATION NUMBER: US/09/296,662B  
CURRENT FILING DATE: 1999-04-22  
PRIOR APPLICATION NUMBER: PCT/US99/08774  
PRIOR FILING DATE: 1999-04-22  
PRIOR APPLICATION NUMBER: 60/000,060  
PRIOR FILING DATE: 1998-04-22

NUMBER OF SEQ ID NOS: 37  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 32  
LENGTH: 2101  
TYPE: PRT  
ORGANISM: Human  
US-09-296-662B-32

Query Match 55.4%; Score 31; DB 5; Length 2101;  
Best Local Similarity 45.5%; Pred. No. 6.8e+02;  
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALVQQMEQLRQ 12  
Db 1075 AQIKELBELRQ 1085

ND

RESULT 40  
US-09-296-662B-33  
Sequence 33, Application US/09296662B  
GENERAL INFORMATION:  
APPLICANT: Anthony Rosen  
APPLICANT: Donald W. Nicholson  
APPLICANT: Livia Casciola-Rosen  
APPLICANT: Felipe A. Andrade  
APPLICANT: Sophie Roy  
APPLICANT: Nancy A. Thornberry  
TITLE OF INVENTION: AUTOANTIGENIC FRAGMENTS, METHODS AND  
TITLE OF INVENTION: ASSAYS  
FILE REFERENCE: 20221Y  
CURRENT APPLICATION NUMBER: US/09/296,662B  
CURRENT FILING DATE: 1999-04-22  
PRIOR APPLICATION NUMBER: PCT/US99/08774  
PRIOR FILING DATE: 1999-04-22  
PRIOR APPLICATION NUMBER: 60/000,060  
PRIOR FILING DATE: 1998-04-22  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 33  
LENGTH: 2115  
TYPE: PRT  
ORGANISM: Human  
US-09-296-662B-33

Query Match 55.4%; Score 31; DB 5; Length 2115;  
Best Local Similarity 45.5%; Pred. No. 6.9e+02;  
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALVQQMEQLRQ 12  
Db 1075 AQIKELBELRQ 1085

ND

Search completed: September 8, 2003, 16:10:45  
Job time : 9 secs



RESULT	1
AAR39483	
ID	AAR39483 standard; Protein; 194 AA.
XX	
AC	AAR39483;
DT	25-MAR-2003 (updated)
DT	28-JAN-1994 (first entry)
XX	
DE	Human apoAIV mutein P(deltaI182).
XX	
KW	apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis
KM	coronary disease; chylomicron; cholesterol transport;
XX	plaque deposition; lecithin-cholesterol acyltransferase; LCAT.
OS	Synthetic.
FH	
FT	Key Location/Qualifiers
FT	Protein 1..194
FT	/label= P(deltaI182)
FT	/note= "the 182 N-terminal amino acids of
XX	mature human apoAIV have been deleted"
XX	
XX	MO9315198-A1.
XX	
PD	05-AUG-1993.
XX	
PF	26-JAN-1993; 93WC-FR00073.
XX	
PR	27-JAN-1992; 92FR-000806.
XX	
PA	(RHON ) RHONE POULENC ROBER SA

XX Denefle P, Guinet F, Latra M, Murry-Brelrier A;  
XX  
XX WPI; 1993-258676/32.  
DR  
XX New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug  
PT design and/or treatment of hypercholesterolaemia  
PT  
XX  
PS Claim 7; Page 31-32 and Page 4; 42pp; French.

Human apolipoprotein AIV is a major component of chylomicrons in lymph but is mainly found in the plasma in unassociated form. The apolip mediates inverse transport of cholesterol. The invention covers polypeptides which are derived from apoAIV by deletion of at least 10 terminal amino acids, by deletion of a helix or pair of helices, by addition of a heterologous polypeptide portion or by a point mutation. Such mutants are useful in the design of hypocholesterolaemic drugs to treat hypercholesterolaemia and atherosclerosis. AAR39483 is a specifically claimed muttein and the sequence has been compiled from the wild-type sequence (see AAR39443) and the description given in the disclosure; the mutant sequence is not shown in the specification.  
(Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 194 AA;

Query Match 100.0%; Score 56; DB 14; Length 194;  
Best Local Similarity 100.0%; Pred. NO. 0.057;  
Matches 12 Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KALVOOMEOLRO 12  
IIIIIIIIIIIIIIIIIIII  
IIIIIIIIIIIIIIIIIIII  
114 KALVOOMEOLRO 125

Db

RESULT 2  
ID AAR39484 Lys 312 from Val Glu Glu Met Glu  
AAR39484 standard; Protein; 328 AA.

XX AC AAR39484;  
XX AC  
XX 25-MAR-2003 (updated)  
DT 28-JAN-1994 (first entry)

XX Human apoAIV muttin P(delta h1-2).  
XX  
XX apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
KM coronary disease; chylomicron; cholesterol transport;  
KW plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.  
XX  
OS Synthetic.

XX  
XX Key Location/Qualifiers  
FH Misc-difference 13..14  
FT /note= "helices 1-2 have been deleted from between  
FT these two sites"  
FT Region 14..46  
FT /note= "helix 3"  
FT Region 47..68  
FT /note= "helix 4"  
FT Region 69..90  
FT /note= "helix 5"  
FT Region 91..112  
FT /note= "helix 6"  
FT Region 113..134  
FT /note= "helix 7"  
FT Region 135..156  
FT /note= "helix 8"  
FT Region 157..178  
FT /note= "helix 9"  
FT Region 179..200  
FT /note= "helix 10"  
FT Region 201..218

FT		/note= "helix 11"	
FT	Region	219..240	
FT		/note= "helix 12"	
FT	Region	241..262	
FT		/note= "helix 13"	
FT	Region	263..284	
FT		/note= "helix 14"	
FT	Region	285..328	
FT		/note= "helix 15"	
XX			
XX	MO9315198-A1.		
XX	05-AUG-1993.		
PD			
PP	26-JAN-1993;	93MO-FR00073.	
PR	27-JAN-1992;	92FR-0000806.	
PR	(RHON ) RHONE POULENC RORER SA.		
PA			
XX	Denefle P, Guinet F, Latta M, Murry-Brelier A;		
PI	WPI; 1993-258676/32.		
DR			
XX			
PT	New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug		
PT	design and/or treatment of hypercholesterolaemia		
PS	Claim 7; Page 31-32 and Page 4; 42pp; French.		
XX			
CC	Human apolipoprotein AIV is a major component of chylomicrons in		
CC	lymph but is mainly found in the plasma in unassociated form. The		
CC	apoAIV mediates inverse transport of cholesterol. The invention		
CC	covers polypeptides which are derived from apoAIV by deletion of		
CC	at least 10 terminal amino acids, by deletion of a helix or pair		
CC	of helices, by addition of a heterologous polypeptide portion or		
CC	by a point mutation. Such muteins are useful in the design of		
CC	hypercholesterolemic drugs to treat hypercholesterolaemia and		
CC	atherosclerosis. AAR39484 is a specifically claimed mutin and the		
CC	sequence has been compiled from the wild-type sequence (see AAR3943)		
CC	and the description given in the disclosure; the mutant sequence is		
CC	not shown in the specification.		
CC	(Updated on 25-MAR-2003 to correct PN field.)		
XX			
SQ	Sequence 328 AA;		
	Query Match	100.0%;	Score 56; DB 14; Length 328;
	Best Local Similarity	100.0%;	Pred. No. 0.097;
	Matches 12; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
OY	1 KALVQMEOQLRK 12		
DB	248 KALVQMEOQLRK 259		
	*		
RESULT 3			
AAR39481			
ID	AAR39481 standard; Protein; 333 AA.		
XX			
AC	AAR39481;		
DT	25-MAR-2003 (updated)	*	
DT	28-JAN-1994 (first entry)		
XX			
DE	Human apoAIV mutuin P(deltaC44).		
XX			
KM	apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;		
KW	coronary disease; chylomicron; cholesterol transport;		
KW	plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.		
OS	Synthetic.		
XX			
FH	Key Location/Qualifiers		
FT	Protein 2..333		

```

FT      /label= P(deliAc44)
FT      /note= "the 44 C-terminal amino acids of
FT      human apoAIV have been detected"
XX
XX      MO9315198-A1-
XX      05-AUG-1993.
XX
XX      26-JAN-1993; 93WO-FR00073.
XX      27-JAN-1992; 92FR-0000806.
XX      PR
XX      (RHON ) RHONE POULENC RORER SA.
XX      Denefle P, Guinet F, Latra M, Murry-Brelier A;
XX      WPI; 1993-258676/32.
XX      DR
XX      New apo:1ipoprotein AIV mutant polypeptide(s) - useful for drug
PT      design and/or treatment of hypercholesterolaemia
XX      Claim 7; Page 31-32 and Page 4; 42pp; French.
XX      PS
XX      Human apolipoprotein AIV is a major component of chylomicrons in
CC      lymph but is mainly found in the plasma in unassociated form. The
CC      apoAIV mediates inverse transport of cholesterol. The invention
CC      covers polypeptides which are derived from apoAIV by deletion of
CC      at least 10 terminal amino acids, by deletion of a helix or pair
CC      of helices, by addition of a heterologous polypeptide portion or
CC      by a point mutation. Such mutants are useful in the design of
CC      hypercholesterolaemic drugs to treat hypercholesterolaemia and
CC      atherosclerosis. AAR39481 is a specifically claimed mutein and the
CC      sequence has been compiled from the wild-type sequence (see AAR39443)
CC      and the description given in the disclosure; the mutant sequence is
CC      not shown in the specification.
CC      (Updated on 25-MAR-2003 to correct PN field.)
XX      SO
XX      Sequence 333 AA;
XX
XX      Query Match 100.0%; Score 56; DB 14; Length 333;
XX      Best Local Similarity 100.0%; Pred. No. 0.098;
XX      Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 KALVQOMEQLRQ 12
DB      297 KALVQOMEQLRQ 308
XX
XX      RESULT 4
XX      ID AAR39488
XX      AAR39488 standard; Protein; 333 AA.
XX      AAR39488;
XX      AC
XX      25-MAR-2003 (updated)
XX      DT 28-JAN-1994 (first entry)
XX      XX
XX      Human apoAIV mutein P(delta h7-8).
XX      XX
XX      apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;
XX      coronary disease; chylomicron; cholesterol transport;
XX      KW plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.
XX      XX
XX      OS Synthetic.
XX      XX
XX      Key Location/Qualifiers
XX      FH 14..40
XX      FT /note= "helix 1"
XX      FT 41..62
XX      FT /note= "helix 2"
XX      FT 63..95
XX      FT /note= "helix 3"
XX      96..117
XX      Region

```

FT	/note= "helix 4"
FT	118..139
FT	/note= "helix 5"
FT	140..161
FT	/note= "helix 6"
FT	161..162
FT	/note= "helices 7-8 have been deleted from between these two sites"
FT	
FT	Region
FT	162..183
FT	/note= "helix 9"
FT	184..205
FT	/note= "helix 10"
FT	206..223
FT	/note= "helix 11"
FT	224..245
FT	/note= "helix 12"
FT	246..267
FT	/note= "helix 13"
FT	268..289
FT	/note= "helix 14"
FT	290..333
FT	/note= "helix 15"
FT	
PN	WO9315198-A1
PD	05-AUG-1993.
XX	
XX	26-JAN-1993; 93WO-FR00073.
XX	
PR	27-JAN-1992; 92FR-0000806.
XX	
PA	(RHON ) RHONE POULENC RORER SA.
XX	
P1	Denefle P, Guinet F, Latta M, Murry-Brelrier A;
XX	
XX	WPI; 1993-258676/32.
XX	
PT	New apo.lipoprotein AIv mutant polypeptide(s) - useful for drug
PT	design and/or treatment of hypercholesterolaemia
XX	
PS	Claim 7; Page 31-32 and Page 4; 42pb; French.
XX	
CC	Human apolipoprotein ATV is a major component of chylomicrons in
CC	lymph but is mainly found in the plasma in unassociated form. The
CC	apoAIV mediates inverse transport of cholesterol. The invention
CC	covers polypeptides which are derived from apoAIV by deletion of
CC	at least 10 terminal amino acids, by deletion of a helix or pair
CC	of helices, by addition of a heterologous polypeptide portion or
CC	by a point mutation. Such muteins are useful in the design of
CC	hypercholesterolemic drugs to treat hypercholesterolaemia and
CC	atherosclerosis. AA939488 is a specifically claimed muten and the
CC	sequence has been compiled from the wild-type sequence (see AA939443)
CC	and the description given in the disclosure; the mutant sequence is
CC	not shown in the specification.
CC	(Updated on 25-MAR-2003 to correct PN field.)
SQ	Sequence 333 AA;
OY	Query Match 100.0%; Score 56; DB 14; Length 333;
	Best Local Similarity 100.0%; Pred. No. 0.098;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
	1 KALVOOMEOLRO 12
DB	253 KALVOOMEOLRO 264
RESULT 5	
AAR39490	
ID	AAR39490 standard; Protein; 333 AA.
XX	
XX	AAR39490;

DT 25-MAR-2003 (updated)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apoAIV mutein P(delta h9-10).  
 KM apolipoprotein AIV, apoAIV, hypercholesterolaemia; atherosclerosis;  
 KM coronary disease; chylomicron; cholesterol transport;  
 KM plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 14..40  
 FT /note= "helix 1"  
 FT Region 41..62  
 FT /note= "helix 2"  
 FT Region 63..95  
 FT /note= "helix 3"  
 FT Region 96..117  
 FT /note= "helix 4"  
 FT Region 118..139  
 FT /note= "helix 5"  
 FT Region 140..161  
 FT /note= "helix 6"  
 FT Region 162..183  
 FT /note= "helix 7"  
 FT Region 184..205  
 FT /note= "helix 8"  
 FT Misc-difference 205..206  
 FT /note= "helices 9-10 have been deleted from between these two sites"  
 FT Region 206..223  
 FT /note= "helix 11"  
 FT Region 224..245  
 FT /note= "helix 12"  
 FT Region 246..267  
 FT /note= "helix 13"  
 FT Region 268..289  
 FT /note= "helix 14"  
 FT Region 290..333  
 FT /note= "helix 15"  
 XX  
 PN WO9315198-A1.  
 XX  
 PD 05-AUG-1993.  
 XX  
 PF 26-JAN-1993; 93WO-FR00073.  
 XX  
 PR 27-JAN-1992; 92FR-000806.  
 XX  
 PA (RHON ) RHONE POULENC RORER SA.  
 XX  
 PI Denefle P, Guinet F, Latta M, Murry-Brellier A;  
 XX WPI; 1993-258676/32.  
 DR  
 XX  
 PT New apo.lipoprotein AIV mutant polypeptide(s) - useful for drug  
 PT design and/or treatment of hypercholesterolaemia  
 XX  
 PS Claim 7; Page 31-32 and Page 5; 42pp; French.  
 XX  
 CC Human apolipoprotein AIV is a major component of chylomicrons in  
 CC lymph but is mainly found in the plasma in unassociated form. The  
 CC apoAIV mediates inverse transport of cholesterol. The invention  
 CC covers polypeptides which are derived from apoAIV by deletion of  
 CC at least 10 terminal amino acids, by deletion of a helix or pair  
 CC of helices, by addition of a heterologous polypeptide portion or  
 CC by a point mutation. Such muteins are useful in the design of  
 CC hypocholesterolaemic drugs to treat hypercholesterolaemia and  
 CC atherosclerosis. AAR39490 is a specifically claimed mutein and the  
 CC sequence has been compiled from the wild-type sequence (see AAR39443)  
 CC and the description given in the disclosure; the mutant sequence is  
 CC not shown in the specification.

CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 333 AA;  
 Query Match 100.0%; Score 56; DB 14; Length 333;  
 Best Local Similarity 100.0%; Pred. No. 0.098;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KALVQOMEQLRQ 12  
 DB 253 KALVQOMEQLRQ 264  
 RESULT 6  
 AAR39497  
 ID AAR39497 standard; Protein; 333 AA.  
 XX  
 AC AAR39497;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apoAIV mutein P(delta h5-6).  
 XX  
 KM apolipoprotein AIV, apoAIV, hypercholesterolaemia; atherosclerosis;  
 KM coronary disease; chylomicron; cholesterol transport;  
 KM plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 14..40  
 FT /note= "helix 1"  
 FT Region 41..62  
 FT /note= "helix 2"  
 FT Region 63..95  
 FT /note= "helix 3"  
 FT Region 96..117  
 FT /note= "helix 4"  
 FT Misc-difference 117..118  
 FT /note= "helices 5-6 have been deleted from between these two sites"  
 FT Region 118..139  
 FT /note= "helix 7"  
 FT Region 140..161  
 FT /note= "helix 8"  
 FT Region 162..183  
 FT /note= "helix 9"  
 FT Region 184..205  
 FT /note= "helix 10"  
 FT Region 206..223  
 FT /note= "helix 11"  
 FT Region 224..245  
 FT /note= "helix 12"  
 FT Region 246..267  
 FT /note= "helix 13"  
 FT Region 268..289  
 FT /note= "helix 14"  
 FT Region 290..333  
 FT /note= "helix 15"  
 XX  
 PN WO9315198-A1.  
 XX  
 PD 05-AUG-1993.  
 XX  
 PF 26-JAN-1993; 93WO-FR00073;  
 XX  
 PR 27-JAN-1992; 92FR-000806.  
 XX  
 PA (RHON ) RHONE POULENC RORER SA.  
 XX  
 PI Denefle P, Guinet F, Latta M, Murry-Brellier A;  
 XX

DR		MPI; 1993-258676/32.
XX		
PT	New apo-IIipoprotein AIV mutant polypeptide(s) - useful for drug design and/or treatment of hypercholesterolaemia	
FT		
XX		
PS	Claim 7; Page 31-32 and Page 5; 42pp; French.	
XX		
CC	Human apolipoprotein AIV is a major component of chylomicrons in lymph but is mainly found in the plasma in unassociated form. The apolip mediates inverse transport of cholesterol. The invention covers polypeptides which are derived from apoAIV by deletion of at least 10 terminal amino acids, by deletion of a helix or pair of helices, by addition of a heterologous polypeptide portion or by a point mutation. Such mutants are useful in the design of hypocholesterolaeamic drugs to treat hypercholesterolaemia and atherosclerosis. AAR39497 is a specifically claimed mutin and the sequence has been compiled from the wild-type sequence (see AAR39443) and the description given in the disclosure; the mutant sequence is not shown in the specification.	
CC	(Updated on 25-MAR-2003 to correct PN field.)	
CC		
XX	Sequence    333 AA;	
SQ		
	Query Match                  100.0%; Score 56; DB 14; Length 333; Best Local Similarity         100.0%; Pred. No. 0.098; Matches      12; Conservative      0; Mismatches    0; Indels    0; Gaps    0	
DY	KALVQOMEQLRQ    12                         KLVQMEQLRQ    264	
Db	253-KLTVQOMELRLQ 264	
RESULT 7		
ID	AAR39485	
AC	AAR39485 standard; Protein; 337 AA.	
XX		
DT	25-MAR-2003 (updated)	
DT	28-JAN-1994 (first entry)	
XX		
DE	Human apoAIV mutein P(tag-delta h1-2).	
XX		
KM	apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis; coronary disease; chylomicron; cholesterol transport;	
KW	plaque deposition; lecithin-cholesterol-acyltransferase; LCAT;	
XX	tag purification marker.	
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	/label= "tag decapeptide"
FT		/note= "allows rapid, single stage purification without affecting the properties of the mature mutelin"
FT	Misc-difference	/note= "helices 1-2 have been deleted from between these two sites"
FT		
FT	Region	23..55 /note= "helix 3"
FT	Region	56..77 /note= "helix 4"
FT	Region	78..99 /note= "helix 5"
FT	Region	100..121 /note= "helix 6"
FT	Region	122..143 /note= "helix 7"
FT	Region	145..165 /note= "helix 8"
FT	Region	166..187 /note= "helix 9"

FT Region 188..209 /note= "helix 10"  
FT FT 210..227 /note= "helix 11"  
FT Region 228..249 /note= "helix 12"  
FT FT 250..271 /note= "helix 13"  
FT Region 272..293 /note= "helix 14"  
FT FT 294..337 /note= "helix 15"  
FT Region  
FN  
PN WO3J15198-A1. *BALME*  
PD 05-AUG-1993.  
PP 26-JAN-1993; 93WO-FR00073.  
PR 27-JAN-1992; 92FR-0000806.  
PS (RHON ) RHONE POULENC RORER SA.  
PI Denefle P, Guinet F, Latra M, Murry-Brelier A;  
PX WPI; 1993-258676/32.  
XX  
DR New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug  
XX design and/or treatment of hypercholesterolaemia  
PT Claim 7; Page 31-32 and Page 4; 42pp; French.  
PS  
XX Human apolipoprotein AIV is a major component of chylomicrons in  
CC lymph but is mainly found in the plasma in unassociated form. The  
CC apoAIV mediates inverse transport of cholesterol. The invention  
CC covers polypeptides which are derived from apoAIV by deletion of  
CC at least 10 terminal amino acids, by deletion of a helix or pair  
CC of helices, by addition of a heterogeneous polypeptide portion or  
CC by a point mutation. Such mutants are useful in the design of  
CC hypocholesterolemic drugs to treat hypercholesterolaemia and  
CC atherosclerosis. AA39485 is a specifically claimed mutein and the  
CC sequence has been compiled from the wild-type sequence (see AA39443)  
CC and the description given in the disclosure; the mutant sequence is  
CC not shown in the specification.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
SQ Sequence 337 AA;

Query Match 100.0%; Score 56; DB 14; Length 337;  
Best Local Similarity 100.0%; Prid. No. 0.099;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

OY 1 KALVQOMEQRLQ 12  
DB 257 KALVQOMEQRLQ 268

RESULT 8  
ID AAR39492 standard; Protein: 337 AA.  
XX AAR39492;  
AC  
XX 25-MAR-2003 (updated)  
DT 26-JAN-1994 (first entry)  
XX  
DE Human apoAIV mutein P(delta, h11-12).  
XX  
KW Apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
KM coronary disease; chylomicron; cholesterol transport;  
KS plaque deposition; lecithin-cholesterol acyltransferase; LCAT.  
SS Synthetic.

XX	Key	Location/Qualifiers
FH	Region	14..40
FT		/note= "helix 1"
FT	Region	41..62
FT		/note= "helix 2"
FT	Region	63..95
FT		/note= "helix 3"
FT	Region	96..117
FT		/note= "helix 4"
FT	Region	118..139
FT		/note= "helix 5"
FT	Region	140..161
FT		/note= "helix 6"
FT	Region	162..183
FT		/note= "helix 7"
FT	Region	184..205
FT		/note= "helix 8"
FT	Region	206..227
FT		/note= "helix 9"
FT	Region	228..249
FT		/note= "helix 10"
FT	Misc-difference	249..250
FT		/note= "helices 11-12 have been deleted from between these two sites"
FT	Region	250..271
FT		/note= "helix 13"
FT	Region	272..293
FT		/note= "helix 14"
FT	Region	294..337
FT		/note= "helix 15"
XX	WO9315198-A1.	
XX	05-AUG-1993.	
XX	26-JAN-1993;	93WO-FR00073.
XX	27-JAN-1992;	92FR-0000806.
XX	(RHON ) RHONE POULENC RORER SA.	
XX	Deneffe P, Guinet F, Latra M, Murry-Brelier A;	
XX	WPI; 1993-258676/32.	
XX	New apo:11ipoprotein AIV mutant polypeptide(s) - useful for drug design and/or treatment of hypercholesterolaemia	
XX	Claim 7; Page 31-32 and Page 5; 42pp; French.	
XX	Human apolipoprotein AIV is a major component of chylomicrons in lymph but is mainly found in the plasma in unassociated form. The apolipoprotein mediates inverse transport of cholesterol. The invention covers polypeptides which are derived from apolipoprotein AIV at least 10 terminal amino acids, by deletion of a helix or pair of helices, by addition of a heterologous polypeptide portion or by a point mutation. Such mutants are useful in the design of hypocholesterolaemic drugs to treat hypercholesterolaemia and atherosclerosis. AAR39492 is a specifically claimed mutant and the sequence has been compiled from the wild-type sequence (see AAR39443) and the description given in the disclosure; the mutant sequence is not shown in the specification.	
XX	(Updated on 25-MAR-2003 to correct PN field.)	
XX	Sequence	337 AA;
XX	Query Match	100.0%; Score 56; DB 14; Length 337;
XX	Best Local Similarity	100.0%; Pred. No. 0.099;
XX	Matches 12; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
XX	1 KALVQOMELRQ 12	

Db	257	KALVQOMEGLRG	268
RESULT	9		
XX	AAR39494		
XX	AAR39494	standard; Protein; 317 AA.	
XX	AAR39494;		
XX	AC		
XX	25-MAR-2003	(updated)	
XX	DT	28-JAN-1994	(first entry)
XX	XX		
DE	Human apoAIV mutein P(delta h11-12, L87M).		
XX	apolipoprotein AIV; apoAIV; hypercholesterolemia; atherosclerosis;		
XX	coronary disease; chylomicron; cholesterol transport;		
KW	plaque deposition; lecithin-cholesterol acyltransferase; LCAT.		
XX	XX		
OS	Synthetic.		
XX	Key	Location/Qualifiers	
FH	Protein	2..317	
FT		/label= (delta h11-12, L87M)	
FT		/note= "helices 11-12 have been deleted from	
FT		mature human apoAIV and wild-type Leu87	
FT		has been substd. by Met"	
FT	Region	14..40	
FT		/note= "helix 1"	
FT	Region	41..62	
FT		/note= "helix 2"	
FT	Region	63..95	
FT		/note= "helix 3"	
FT	Misc-difference	87	
FT		/note= "L87M"	
FT	Region	96..117	
FT		/note= "helix 4"	
FT	Region	118..139	
FT		/note= "helix 5"	
FT	Region	140..161	
FT		/note= "helix 6"	
FT	Region	162..183	
FT		/note= "helix 7"	
FT	Region	184..205	
FT		/note= "helix 8"	
FT	Region	206..227	
FT		/note= "helix 9"	
FT	Region	228..249	
FT		/note= "helix 10"	
FT	Misc-difference	249..250	
FT		/note= "helices 11-12 have been deleted from between	
FT		these two sites"	
FT	Region	250..271	
FT		/note= "helix 13"	
FT	Region	272..293	
FT		/note= "helix 14"	
FT	Region	294..337	
FT		/note= "helix 15"	
XX	XX		
XX	PN	WO9315198-A1.	
XX	PD	05-AUG-1993.	
XX	XX		
XX	PF	26-JAN-1993;	93WO-FR00073.
XX	XX		
XX	PR	27-JAN-1992;	92FR-0000806.
XX	XX		
XX	PA	(RHON ) RHONE POULENC ROBER SA.	
XX	PI		
XX	Denefle P, Guinet F, Latte M, Murry-Brelrier A;		
XX	DR	WPI; 1993-258676/32.	
XX	XX		
PT	New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug		

PT design and/or treatment of hypercholesterolaemia  
 XX  
 PS Claim 7; Page 31-32 and Page 5; 42pp; French.  
 XX  
 CC Human apolipoprotein AIV is a major component of chylomicrons in  
 CC lymph but is mainly found in the plasma in unassociated form. The  
 CC apoAIV mediates inverse transport of cholesterol. The invention  
 CC covers polypeptides which are derived from apoAIV by deletion of  
 CC at least 10 terminal amino acids, by deletion of a helix or pair  
 CC of helices, by addition of a heterologous polypeptide portion or  
 CC by a point mutation. Such mutants are useful in the design of  
 CC hypocholesterolaemic drugs to treat hypercholesterolaemia and  
 CC atherosclerosis. AAR39487 is a specifically claimed mutein and the  
 CC sequence has been compiled from the wild-type sequence (see AAR39443)  
 CC and the description given in the disclosure; the mutant sequence is  
 CC not shown in the specification.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 XX  
 SQ Sequence 337 AA;  
 Query Match 100.0%; Score 56; DB 14; Length 337;  
 Best Local Similarity 100.0%; Pred. NO. 0.099;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KALVQOMEQLRQ 12  
 Db 257 KALVQOMEQLRQ 268  
 RESULT 10  
 AAR39487  
 ID AAR39487 standard; Protein; 342 AA.  
 XX  
 AC AAR39487;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apoAIV mutein P(tag-deltaC44).  
 XX  
 KW apolipoprotein AIV, apoAIV, hypercholesterolaemia; atherosclerosis;  
 KW coronary disease; chylomicron; cholesterol transport;  
 KW plaque deposition; lecithin-cholesterol-acyltransferase; LCAT;  
 KW tag purification marker.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FH Peptide 1..10  
 FT /label= tag-decapeptide  
 FT /note= "allows rapid, single stage purification  
 FT without affecting the properties of the  
 FT mature mutein"  
 FT Protein 11..342  
 FT /label= P(deltaC44)  
 FT /note= "the 44 C-terminal amino acids of  
 FT mature human apoAIV have been deleted"  
 XX  
 PN W09315198-A1.  
 XX  
 PD 05-AUG-1993.  
 XX  
 PD 26-JAN-1993; 93WO-FR00073.  
 XX  
 PF 27-JAN-1992; 92FR-0000806.  
 XX  
 PR (RHON ) RHONE POTLENC RORER SA.  
 PA  
 PI Deneffe P, Guinet F, Latra M, Murry-Brelier A;  
 XX WPI; 1993-258676/32.  
 DR  
 PT New apo:11poprotein AIV mutant polypeptide(s) - useful for drug

PT design and/or treatment of hypercholesterolaemia  
 XX  
 PS Claim 7; Page 31-32 and Page 4; 42pp; French.  
 XX  
 CC Human apolipoprotein AIV is a major component of chylomicrons in  
 CC lymph but is mainly found in the plasma in unassociated form. The  
 CC apoAIV mediates inverse transport of cholesterol. The invention  
 CC covers polypeptides which are derived from apoAIV by deletion of  
 CC at least 10 terminal amino acids, by deletion of a helix or pair  
 CC of helices, by addition of a heterologous polypeptide portion or  
 CC by a point mutation. Such mutants are useful in the design of  
 CC hypocholesterolaemic drugs to treat hypercholesterolaemia and  
 CC atherosclerosis. AAR39487 is a specifically claimed mutein and the  
 CC sequence has been compiled from the wild-type sequence (see AAR39443)  
 CC and the description given in the disclosure; the mutant sequence is  
 CC not shown in the specification.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 XX  
 SQ Sequence 342 AA;  
 Query Match 100.0%; Score 56; DB 14; Length 342;  
 Best Local Similarity 100.0%; Pred. NO. 0.1; 0;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KALVQOMEQLRQ 12  
 Db 306 KALVQOMEQLRQ 317  
 RESULT 11  
 AAR39489  
 ID AAR39489 standard; Protein; 342 AA.  
 XX  
 AC AAR39489;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apoAIV mutein P(tag-delta h7-8).  
 XX  
 KW apolipoprotein AIV, apoAIV, hypercholesterolaemia; atherosclerosis;  
 KW coronary disease; chylomicron; cholesterol transport;  
 KW plaque deposition; lecithin-cholesterol-acyltransferase; LCAT;  
 KW tag purification marker.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FH Peptide 1..10  
 FT /label= tag-decapeptide  
 FT /note= "allows rapid, single stage purification  
 FT without affecting the properties of the  
 FT mature mutein"  
 FT Region 23..49  
 FT /note= "helix 1"  
 FT Region 50..71  
 FT /note= "helix 2"  
 FT Region 72..104  
 FT /note= "helix 3"  
 FT Region 105..126  
 FT /note= "helix 4"  
 FT Region 127..148  
 FT /note= "helix 5"  
 FT Region 149..170  
 FT /note= "helix 6"  
 FT Region 170..171  
 FT /note= "helices 7-8 have been deleted from between  
 FT these two sites"  
 FT Region 171..192  
 FT /note= "helix 9"  
 FT Region 193..214  
 FT /note= "helix 10"  
 FT Region 215..232

FT /note= "helix 11"  
 FT 233..254  
 FT /note= "helix 12"  
 FT 255..276  
 FT /note= "helix 13"  
 FT 277..298  
 FT /note= "helix 14"  
 FT 299..342  
 FT /note= "helix 15"

XX WO9315198-A1.  
 PN  
 PD 05-AUG-1993.

*Sane*

XX 26-JAN-1993; 93WO-FR00073.  
 XX  
 XX 27-JAN-1992; 92FR-0000806.  
 XX  
 XX (RHON ) RHONE POULENC RORER SA.  
 XI Deneŕle P, Guinet F, Latta M, Murry-Brelier A;  
 XX WPI : 1993-258676/32.  
 DR  
 XX  
 PT New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug  
 PT design and/or treatment of hypercholesterolaemia

XX Claim 7 ; Page 31-32 and Page 4 ; 42pp; French.

Human apolipoprotein AIV is a major component of chylomicrons in lymph but is mainly found in the plasma in unassociated form. The apoAIV mediates inverse transport of cholesterol. The invention covers polypeptides which are derived from apoAIV by deletion of at least 10 terminal amino acids, by deletion of a helix or pair of helices, by addition of a heterologous polypeptide portion or by a point mutation. Such muteins are useful in the design of hypocholesterolaemic drugs to treat hypercholesterolaemia and atherosclerosis. AAR39449 is a specifically claimed mutein and the sequence has been compiled from the wild-type sequence (see AAR39443) and the description given in the disclosure; the mutant sequence is not shown in the specification.  
 (Updated on 25-MAR-2003 to correct PN field.)

SO Sequence 342 AA;

Query Match	100.0%;	Score 56;	DB 14;	Length 342;
Best Local Similarity	100.0%;	Pred. No. 0.1;		
Matches 12;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

OY 1 KALVQOMEQLRQ 12  

DB 262 KALVQOMEQLRQ 273

RESULT 12  
 AAR39491  
 ID AAR39491 standard; Protein; 342 AA.  
 XX  
 AC AAR39491;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 28-JAN-1994 (first entry)

XX Human apoAIV mutein P(tag-delta h9-10).  
 DE  
 KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KW coronary disease; chylomicon; cholesterol transport;  
 KW lecithin deposition; lecithin-cholesterol-acyltransferase; LCAT;  
 XX tag purification marker.  
 OS Synthetic.  
 RX Key Location/Qualifiers

```

FT Peptide 1..10
FT /label= rag decapeptide
FT /note= "allows rapid, single stage purification
FT without affecting the properties of the
FT mature muten"
FT Region 23..49
FT /note= "helix 1"
FT Region 50..71
FT /note= "helix 2"
FT Region 72..104
FT /note= "helix 3"
FT Region 105..126
FT /note= "helix 4"
FT Region 127..148
FT /note= "helix 5"
FT Region 149..170
FT /note= "helix 6"
FT Region 171..192
FT /note= "helix 7"
FT Region 193..214
FT /note= "helix 8"
FT Misc-difference 214..215
FT /note= "helices 9-10 have been deleted from between
FT these two sites"
FT Region 215..232
FT /note= "helix 11"
FT Region 233..254
FT /note= "helix 12"
FT Region 255..276
FT /note= "helix 13"
FT Region 277..298
FT /note= "helix 14"
FT Region 299..342
FT /note= "helix 15"
XX
XX WO9315198-A1.
XX
XX PD 05-AUG-1993.
XX
XX PF 26-JAN-1993; 93WO-FR00073.
XX PR 27-JAN-1992; 92FR-0000806.
XX PA (RHON ) RHONE POULENC RORER SA.
XX PI Deneffe P, Guinet F, Latta M, Murry-Brelier A;
XX DR WPI, 1993-258676/32.
XX PT New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug
XX design and/or treatment of hypercholesterolaemia
XX PS Claim 7; Page 31-32 and Page 5; 42pp; French.
XX
CC Human apolipoprotein AIV is a major component of chylomicrons in
CC lymph but is mainly found in the plasma in unassociated form. The
CC apoAIV mediates inverse transport of cholesterol. The invention
CC covers polypeptides which are derived from apoAIV by deletion of
CC at least 10 terminal amino acids, by deletion of a helix or pair
CC of helices, by addition of a heterologous polypeptide portion or
CC by a point mutation. Such muteins are useful in the design of
CC hypocholesterolaemic drugs to treat hypercholesterolemia and
CC atherosclerosis. AA839491 is a specifically claimed mutein and the
CC sequence has been compiled from the wild-type sequence (see AA839443)
CC and the description given in the disclosure; the mutant sequence is
CC not shown in the specification.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 342 AA;
Query Match 100.0%; Score 56; DB 14; Length 342;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



OY 1 KALVQOMEQLRQ 12  
DB 262 KALVQOMEQLRQ 273

RESULT 13  
AAR39498 AAR39498 standard; Protein; 342 AA.  
XX  
AC AAR39498;  
XX  
DT 25-MAR-2003 (updated)  
DT 28-JAN-1994 (first entry)  
XX  
DE Human apoAIV mutein P (tag-delta h5-6).  
XX  
KM apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
KM coronary disease; chylomicron; cholesterol transport;  
KM plaque deposition; lecithin-cholesterol-acyltransferase; LCAT;  
KM tag purification marker.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide  
FT 1..10  
FT /label= tag,decapeptide  
FT /note= "allows rapid, single stage purification  
FT without affecting the properties of the  
FT mature mutein"  
FT  
FT Region 23..49  
FT /note= "helix 1"  
FT Region 50..71  
FT /note= "helix 2"  
FT Region 72..104  
FT /note= "helix 3"  
FT Region 105..126  
FT /note= "helix 4"  
FT Misc-difference 126..127  
FT /note= "helices 5-6 have been deleted from between  
FT these two sites"  
FT  
FT Region 127..148  
FT /note= "helix 7"  
FT Region 149..170  
FT /note= "helix 8"  
FT Region 171..192  
FT /note= "helix 9"  
FT Region 193..214  
FT /note= "helix 10"  
FT Region 215..232  
FT /note= "helix 11"  
FT Region 233..254  
FT /note= "helix 12"  
FT Region 255..276  
FT /note= "helix 13"  
FT Region 277..298  
FT /note= "helix 14"  
FT Region 299..342  
FT /note= "helix 15"  
XX  
XX WO9315198-A1.  
XX  
XX 05-AUG-1993.  
XX  
XX 26-JAN-1993; 93WO-FR00073.  
XX  
XX 27-JAN-1992; 92FR-0000806.  
XX  
XX (RHON ) RHONE POULENC RORER SA.  
XX  
XX Deneffe P, Guinet F, Latte M, Murry-Breiller A;  
XX  
XX WPI, 1993-258676/32.  
XX  
XX

XX  
PT New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug  
FT design and/or treatment of hypercholesterolaemia  
XX  
PS Claim 7; Page 31-32 and Page 5; 42pp; French.  
XX  
CC Human apolipoprotein AIV is a major component of chylomicrons in  
CC lymph but is mainly found in the plasma in unassociated form. The  
CC apoAIV mediates inverse transport of cholesterol. The invention  
CC covers polypeptides which are derived from apoAIV by deletion of  
CC at least 10 terminal amino acids, by deletion of a helix or pair  
CC of helices, by addition of a heterologous polypeptide portion or  
CC by a point mutation. Such muteins are useful in the design of  
CC hypocholesterolaemic drugs to treat hypercholesterolaemia and  
CC atherosclerosis. AAR39498 is a specifically claimed mutein and the  
CC sequence has been compiled from the wild-type sequence (see AAR39443)  
CC and the description given in the disclosure; the mutant sequence is  
CC not shown in the specification.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 342 AA:  
XX  
Query Match 100.0%; Score 56; DB 14; Length 342;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 KALVQOMEQLRQ 12  
DB 262 KALVQOMEQLRQ 273  
RESULT 14  
AAR39493 AAR39493 standard; Protein; 346 AA.  
XX  
XX AAR39493;  
XX  
DT 25-MAR-2003 (updated)  
DT 28-JAN-1994 (first entry)  
XX  
XX  
DE Human apoAIV mutein P (tag-delta h11-12).  
XX  
KM apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
KM coronary disease; chylomicron; cholesterol transport;  
KM plaque deposition; lecithin-cholesterol-acyltransferase; LCAT;  
KM tag purification marker.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide  
FT 1..10  
FT /label= tag,decapeptide  
FT /note= "allows rapid, single stage purification  
FT without affecting the properties of the  
FT mature mutein"  
FT  
FT Region 23..49  
FT /note= "helix 1"  
FT Region 50..71  
FT /note= "helix 2"  
FT Region 72..104  
FT /note= "helix 3"  
FT Region 105..126  
FT /note= "helix 4"  
FT Region 127..148  
FT /note= "helix 5"  
FT Region 149..170  
FT /note= "helix 6"  
FT Region 171..192  
FT /note= "helix 7"  
FT Region 193..214  
FT /note= "helix 8"  
FT Region 215..236  
FT /note= "helix 9"  
XX  
XX

FT	Region	237..258	/note= "helix 10"
FT	Misc-difference	258..259	/note= "helices 11-12 have been deleted from between these two sites"
FT	Region	259..280	/note= "helix 13"
FT	Region	281..302	/note= "helix 14"
FT	Region	303..346	/note= "helix 15"
PN	MO9315198-A1.		
XX	05-AUG-1993.		
XX	26-JAN-1993;	93WO-FR00073.	
XX	27-JAN-1992;	92FR-0000806.	
XX	(RHON ) RHONE POULENC RORER SA.		
XX	Denefle P, Guinet F, Latra M, Murry-Brelier A;		
XX	WPI, 1993-258676/32.		
PT	New apo:1ipoprotein AIV mutant polypeptide(s) - useful for drug design and/or treatment of hypercholesterolaemia		
PS	Claim 7; Page 31-32 and Page 5; 42pp; French.		
XX	Human apolipoprotein AIV is a major component of chylomicrons in lymph but is mainly found in the plasma in unassociated form. The apoAIV mediates inverse transport of cholesterol. The invention covers polypeptides which are derived from apoAIV by deletion of at least 10 terminal amino acids, by deletion of a helix or pair of helices, by addition of a heterologous polypeptide portion or by a point mutation. Such mutants are useful in the design of hypocholesterolaemic drugs to treat hypercholesterolaemia and atherosclerosis. AAR39493 is a specifically claimed mutant and the sequence has been compiled from the wild-type sequence (see AAR39443) and the description given in the disclosure; the mutant sequence is not shown in the specification.		
XX	(Updated on 25-MAR-2003 to correct PN field.)		
SQ	Sequence	346 AA;	
	Query Match	100.0%; Score 56; DB 14; Length 346;	
	Best Local Similarity	100.0%; Pred. No. 0.1;	
	Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1 KALVQOMEQLRQ 12		
DB	266 KALVQOMEQLRQ 277		
RESULT 15			
ID	AAR39478	standard; Protein; 363 AA.	
XX	AAR39478;		
XX	25-MAR-2003 (updated)		
DT	28-JAN-1994 (first entry)		
XX	Human apoAIV mutein P(delta1N13, R93G) .		
XX	apolipoprotein AIV; apoAIV: hypercholesterolaemia; atherosclerosis; coronary disease; chylomicron; cholesterol transport; plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.		
XX	Synthetic.		

Key	Location/Qualifiers
FT	1..363
FT	/label= "deltaN13, R93G"
FT	/note= "the 13 N-terminal amino acids of human apoAIV have been deleted and wild-type Arg93 has been subst. by Gly"
FT	Misc-difference 79
FT	/note= "R93G"
XX	
PN	W09315198-A1.
XX	
PD	05-AUG-1993.
XX	
PF	26-JAN-1993; 93MO-FR00073.
XX	
PR	27-JAN-1992; 92FR-0000806.
XX	
PA	(RHON ) RHONE POULENC RORER SA.
PI	Deneffe P, Guinet F, Latta M, Murry-Brelier A;
DR	WPI; 1993-258676/32.
XX	
PT	New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design and/or treatment of hypercholesterolaemia
XX	
PS	Claim 7; Page 31-32 and Page 4; 42pp; French.
XX	
CC	Human apolipoprotein AIV is a major component of chylomicrons in lymph but is mainly found in the plasma in unassociated form. The apoAIV mediates inverse transport of cholesterol. The invention covers polypeptides which are derived from apoAIV by deletion of at least 10 terminal amino acids, by deletion of a helix or pair of helices, by addition of a heterologous polypeptide portion or by a point mutation. Such mutants are useful in the design of hypocholesterolemic drugs to treat hypercholesterolaemia and atherosclerosis. AAR39478 is a specifically claimed muten and the sequence has been compiled from the wild-type sequence (see AAR39443) and the description given in the disclosure; the mutant sequence is not shown in the specification.
CC	(Updated on 25-MAR-2003 to correct PN field.)
XX	
SO	Sequence 363 AA;
QY	Query Match 100.0%; Score 56; DB 14; Length 363; Best Local Similarity 100.0%; Prid. NO. 0.11; Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	1 KALVOQMEQLRQ 12                         283 KALVOQMEQLRQ 294
RESULT 16	
ID	AAR39479 standard; Protein; 363 AA.
XX	
AC	AAR39479;
XX	
DT	25-MAR-2003 (updated)
DT	28-JAN-1994 (first entry)
XX	
DE	Human apoAIV muten P(deltaN13).
XX	
KW	apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis; coronary disease; chylomicron; cholesterol transport; plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.
XX	
OS	Synthetic.
XX	
Key	Location/Qualifiers
FT	1..363
FT	/label= P(deltaN13)

FT /note= "the 13 N-terminal amino acids of  
 FT mature human apoAIV have been deleted"  
 XX MO9315198-A1. *Shw*  
 XX 05-AUG-1993.  
 PD  
 XX 26-JAN-1993; 93WO-FR00073.  
 XX 27-JAN-1992; 92FR-0000806.  
 XX (RHON ) RHONE POULENC RORER SA.  
 XX Deneffe P, Guinet F, Latta M, Murry-Brelier A;  
 XX WPI; 1993-258676/32.  
 DR  
 XX  
 PT New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug  
 PT design and/or treatment of hypercholesterolaemia  
 XX  
 PS Claim 7; Page 31-32 and Page 4; 42pp; French.  
 XX  
 CC Human apolipoprotein AIV is a major component of chylomicrons in  
 CC lymph but is mainly found in the plasma in unassociated form. The  
 CC apoAIV mediates inverse transport of cholesterol. The invention  
 CC covers polypeptides which are derived from apoAIV by deletion of  
 CC at least 10 terminal amino acids, by deletion of a helix or pair  
 CC of helices, by addition of a heterologous polypeptide portion or  
 CC by a point mutation. Such muteins are useful in the design of  
 CC hypocholesterolemic drugs to treat hypercholesterolaemia and  
 CC atherosclerosis. AAR39486 is a specifically claimed mutein and the  
 CC sequence has been compiled from the wild-type sequence (see AAR39443)  
 CC and the description given in the disclosure; the mutant sequence is  
 CC not shown in the specification.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 SQ Sequence 363 AA;  
 Query Match 100.0%; Score 56; DB 14; Length 363;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KALVQOMEQLRQ 12  
 DB 283 KALVQOMEQLRQ 294  
 RESULT 17  
 AAR39486  
 ID AAR39486 standard; Protein; 373 AA.  
 XX  
 AC AAR39486;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apoAIV mutein P (tag-deltaN13).  
 XX  
 KM apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KM coronary disease; chylomicron; cholesterol transport;  
 KW plaque deposition; lecithin-cholesterol-acyltransferase; LCAT;  
 XX tag purification marker.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..10  
 FT /label= tag-decapeptide  
 FT /note= "allows rapid, single stage purification  
 FT without affecting the properties of the  
 FT mature mutein"  
 FT Protein 11..373  
 FT /label= P(deltaN13)

FT /note= "the 13 N-terminal amino acids of  
 FT mature human apoAIV have been deleted"  
 XX MO9315198-A1. *Barnd*  
 XX 05-AUG-1993.  
 PD  
 XX 26-JAN-1993; 93WO-FR00073.  
 XX 27-JAN-1992; 92FR-0000806.  
 XX (RHON ) RHONE POULENC RORER SA.  
 XX Deneffe P, Guinet F, Latta M, Murry-Brelier A;  
 XX WPI; 1993-258676/32.  
 DR  
 XX  
 PT New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug  
 PT design and/or treatment of hypercholesterolaemia  
 XX  
 PS Claim 7; Page 31-32 and Page 4; 42pp; French.  
 XX  
 CC Human apolipoprotein AIV is a major component of chylomicrons in  
 CC lymph but is mainly found in the plasma in unassociated form. The  
 CC apoAIV mediates inverse transport of cholesterol. The invention  
 CC covers polypeptides which are derived from apoAIV by deletion of  
 CC at least 10 terminal amino acids, by deletion of a helix or pair  
 CC of helices, by addition of a heterologous polypeptide portion or  
 CC by a point mutation. Such muteins are useful in the design of  
 CC hypocholesterolemic drugs to treat hypercholesterolaemia and  
 CC atherosclerosis. AAR39486 is a specifically claimed mutein and the  
 CC sequence has been compiled from the wild-type sequence (see AAR39443)  
 CC and the description given in the disclosure; the mutant sequence is  
 CC not shown in the specification.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 SQ Sequence 373 AA;  
 Query Match 100.0%; Score 56; DB 14; Length 373;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KALVQOMEQLRQ 12  
 DB 293 KALVQOMEQLRQ 304  
 RESULT 18  
 AAR39443  
 ID AAR39443 standard; Protein; 377 AA.  
 XX  
 AC AAR39443;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apolipoprotein AIV.  
 XX  
 KM apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KM coronary disease; chylomicron; cholesterol transport;  
 KW plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 14..40  
 FT /note= "helix 1"  
 FT Region 41..62  
 FT /note= "helix 2"  
 FT Region 63..95  
 FT /note= "helix 3"  
 FT Region 96..117  
 FT /note= "helix 4"

FT	Region	118..139
FT	/note="helix 5"	
FT	Region	140..161
FT	/note="helix 6"	
FT	Region	162..183
FT	/note="helix 7"	
FT	Region	184..205
FT	/note="helix 8"	
FT	Region	206..227
FT	/note="helix 9"	
FT	Region	228..249
FT	/note="helix 10"	
FT	Region	250..267
FT	/note="helix 11"	
FT	Region	268..289
FT	/note="helix 12"	
FT	Region	290..311
FT	/note="helix 13"	
FT	Region	312..333
FT	/note="helix 14"	
FT	Region	334..377
FT	/note="helix 15"	
XX	PN	WO315198-A1.
XX	PD	05-AUG-1993.
XX	PF	26-JAN-1993; 93WO-FR00073.
XX	PR	27-JAN-1992; 92FR-0000806.
XX	PA	(RHON ) RHONE POULENC RORER SA.
XX	PI	Deneffe P, Guinet F, Latra M, Murry-Brelier A;
XX	DR	WP1; 1993-258676/32.
XX	DR	N-PSDB; AAQ43306.
PT	PT	New apo:1ipoprotein AIV mutant polypeptide(s) - useful for drug
PT	PT	design and/or treatment of hypercholesterolaemia
XX	XX	
PS	PS	Disclosure; Page 32-33; 42pp; French.
XX	XX	
CC	CC	Human apolipoprotein AIV is a major component of chylomicrons in
CC	CC	lymph but is mainly found in the plasma in unassociated form. The
CC	CC	apo AIV mediates inverse transport of cholesterol. The invention
CC	CC	covers polypeptides which are derived from apoAIV by deletion of
CC	CC	at least 10 terminal amino acids, by deletion of a helix or pair
CC	CC	of helices, by addition of a heterologous polypeptide portion or
CC	CC	by a point mutation. Such mutants are useful in the design of
CC	CC	hypercholesterolaemic drugs to treat hypercholesterolaemia and
CC	CC	atherosclerosis.
CC	CC	(Updated on 25-MAR-2003 to correct PN field.)
XX	XX	
SQ	SQ	Sequence 377 AA;
Query Match	100.0%;	Score 56; DB 14; Length 377;
Best Local Similarity	100.0%;	Pred. NO. 0.11;
Matches 12; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1 KALVQOMEQLRQ 12	
DB	297 KALVQOMEQLRQ 308	
RESULT 19		
ID	AA839480	
XX	AA839480 standard; Protein; 377 AA.	
XX	AA839480;	
XX		
DT	25-MAR-2003 (updated)	
DT	28-JAN-1994 (first entry)	

Human apoAIV mutein P(R93G).

Human apolipoprotein AIV, apoAIV, hypercholesterolaemia, atherosclerosis; coronary disease; chylomicron; cholesterol transport; plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.

Synthetic.

Key Location/Qualifiers  
FT Protein 2..377  
FT /note= "mature P(R93G)"  
FT Misc-difference 94  
FT /note= "wild-type Arg93 is substituted by Gly"

W09315198-A1.

05-AUG-1993.

26-JAN-1993; 93MO-FR00073.

27-JAN-1992; 92FR-0000806.

(RHON ) RHONE POULENC RORER SA.

Denefle P, Guinet F, Lacta M, Murry-Brellier A;  
WPI; 1993-258676/32.

New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design and/or treatment of hypercholesterolaemia

Claim 7; Page 31-32 and Page 4; 42pp; French.

Human apolipoprotein AIV is a major component of chylomicrons in lymph but is mainly found in the plasma in unassociated form. The apoAIV mediates inverse transport of cholesterol. The invention covers polypeptides which are derived from apoAIV by deletion of at least 10 terminal amino acids, by deletion of a helix or pair of helices, by addition of a heterologous polypeptide portion or by a point mutation. Such muteins are useful in the design of hypocholesterolemic drugs to treat hypercholesterolaemia and atherosclerosis. AAR39480 is a specifically claimed mutein and the sequence has been compiled from the wild-type sequence (see AAR39443) and the description given in the disclosure; the mutant sequence is not shown in the specification.  
(Updated on 25-MAR-2003 to correct PN field.)

Sequence 377 AA;

Query Match 100.0%; Score 56; DB 14; Length 377;  
Best Local Similarity 100.0%; Pred. NO. 0.11;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

1 KALVQOMEQLRQ 12  

297 KALVQOMEQLRQ 308

Db

RESULT 20  
AAR39499  
ID AAR39499 standard; Protein; 377 AA.

AC AAR39499;  
XX  
XX 25-MAR-2003 (updated)  
DT 28-JAN-1994 (first entry)  
XX  
XX Human apoAIV mutein P(D44F).  
XX  
XX apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
KW coronary disease; chylomicron; cholesterol transport;  
KW plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.

XX OS Synthetic.  
 XX Key Location/Qualifiers  
 FH Protein 2..377  
 FT /note= "mature P(D44F)"  
 FT Misc-difference 45  
 FT /note= "wild-type Asp44 is substd. by Phe"  
 FT Region 14..40  
 FT /note= "helix 1"  
 FT Region 41..62  
 FT /note= "helix 2"  
 FT Region 63..95  
 FT /note= "helix 3"  
 FT Region 96..117  
 FT /note= "helix 4"  
 FT Region 118..139  
 FT /note= "helix 5"  
 FT Region 140..161  
 FT /note= "helix 6"  
 FT Region 162..183  
 FT /note= "helix 7"  
 FT Region 184..205  
 FT /note= "helix 8"  
 FT Region 206..227  
 FT /note= "helix 9"  
 FT Region 228..249  
 FT /note= "helix 10"  
 FT Region 250..267  
 FT /note= "helix 11"  
 FT Region 268..289  
 FT /note= "helix 12"  
 FT Region 290..311  
 FT /note= "helix 13"  
 FT Region 312..333  
 FT /note= "helix 14"  
 FT Region 334..377  
 FT /note= "helix 15"  
 XX PM W09315198-A1. *Gene*  
 XX PD 05-AUG-1993.  
 XX PF 26-JAN-1993; 93WO-FR00073.  
 XX PR 27-JAN-1992; 92FR-0000806.  
 XX PA (RHON ) RHONE POULENC RORER SA.  
 XX PI Deneffe P, Guinet F, Latta M, Murry-Brellier A;  
 XX DR WPI, 1993-258676/32.  
 XX FT New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug  
 XX PT design and/or treatment of hypercholesterolaemia  
 XX PS Claim 7; Page 31-32 and Page 5; 42pp; French.  
 XX CC Human apolipoprotein AIV is a major component of chylomicrons in  
 CC lymph but is mainly found in the plasma in unassociated form. The  
 CC apo AIV mediates inverse transport of cholesterol. The invention  
 CC covers polypeptides which are derived from apoAIV by deletion of  
 CC at least 10 terminal amino acids, by deletion of a helix or pair  
 CC of helices, by addition of a heterologous polypeptide portion or  
 CC by a point mutation. Such mutants are useful in the design of  
 CC hypocholesterolaemic drugs to treat hypercholesterolaemia and  
 CC atherosclerosis. AAR39499 is a specifically claimed mutant and the  
 CC sequence has been compiled from the wild-type sequence (see AAR39443)  
 CC and the description given in the disclosure; the mutant sequence is  
 CC not shown in the specification.  
 CC (Updated on 25-MAR-2003 to correct PM field.)  
 CC XX Sequence 377 AA;

Query Match 100.0%; Score 56; DB 14; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KALVQMEQLRQ 12  
 DB 297 KALVQMEQLRQ 308  
 RESULT 21  
 AAR39500 standard; Protein; 377 AA.  
 AC AAR39500;  
 XX DT 25-MAR-2003 (updated)  
 DT 28-JAN-1994 (first entry)  
 XX DE Human apoAIV murein P(D44A).  
 XX KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KW coronary disease; chylomicron; cholesterol transport;  
 KW plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.  
 XX OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Protein 2..377  
 FT /note= "mature P(D44A)"  
 FT Misc-difference 45  
 FT /note= "wild-type Asp44 is substd. by Ala"  
 FT Region 14..40  
 FT /note= "helix 1"  
 FT Region 41..62  
 FT /note= "helix 2"  
 FT Region 63..95  
 FT /note= "helix 3"  
 FT Region 96..117  
 FT /note= "helix 4"  
 FT Region 118..139  
 FT /note= "helix 5"  
 FT Region 140..161  
 FT /note= "helix 6"  
 FT Region 162..183  
 FT /note= "helix 7"  
 FT Region 184..205  
 FT /note= "helix 8"  
 FT Region 206..227  
 FT /note= "helix 9"  
 FT Region 228..249  
 FT /note= "helix 10"  
 FT Region 250..267  
 FT /note= "helix 11"  
 FT Region 268..289  
 FT /note= "helix 12"  
 FT Region 290..311  
 FT /note= "helix 13"  
 FT Region 312..333  
 FT /note= "helix 14"  
 FT Region 334..377  
 FT /note= "helix 15"  
 XX PM W09315198-A1. *Gene*  
 XX PD 05-AUG-1993.  
 XX PF 26-JAN-1993; 93WO-FR00073.  
 XX PR 27-JAN-1992; 92FR-0000806.  
 XX PA (RHON ) RHONE POULENC RORER SA.

PI Deneffe P, Guinet F, Latra M, Murry-Brelier A;  
 XX WPI; 1993-258676/32.  
 XX  
 PT New apolipoprotein AIV mutant polypeptide(s) - useful for drug  
 PT design and/or treatment of hypercholesterolaemia  
 XX  
 PS Claim 7; Page 31-32 and Page 5; 42pp; French.  
 XX  
 CC Human apolipoprotein AIV is a major component of chylomicrons in  
 CC lymph but is mainly found in the plasma in unassociated form. The  
 CC apo AIV mediates inverse transport of cholesterol. The invention  
 CC covers polypeptides which are derived from apoAIV by deletion of  
 CC at least 10 terminal amino acids, by deletion of a helix or pair  
 CC of helices, by addition of a heterologous polypeptide portion or  
 CC by a point mutation. Such mutants are useful in the design of  
 CC hypocholesterolaemic drugs to treat hypercholesterolaemia and  
 CC atherosclerosis. AAR39500 is a specifically claimed mutein and the  
 CC sequence has been compiled from the wild-type sequence (see AAR39443)  
 CC and the description given in the disclosure; the mutant sequence is  
 CC not shown in the specification.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 XX  
 SQ Sequence 377 AA;  
 Query Match 100.0%; Score 56; DB 14; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KALVOQMEQLRQ 12  
 DB 297 KALVOQMEQLRQ 308  
 XX  
 RESULT 22  
 AAR39502  
 ID AAR39502 standard; Protein; 377 AA.  
 XX  
 AC AAR39502;  
 DT 25-MAR-2003 (updated)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apoAIV mutein P(D5K).  
 XX  
 KM apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KM coronary disease; chylomicron; cholesterol transport;  
 KM plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT 2..377 /note= "mature P(D5K)"  
 FT Misc-difference 6 /note= "wild-type Asps is substd. by Lys"  
 FT 14..40 /note= "helix 1"  
 FT Region 41..62 /note= "helix 2"  
 FT 63..95 /note= "helix 3"  
 FT Region 96..117 /note= "helix 4"  
 FT 118..139 /note= "helix 5"  
 FT Region 140..161 /note= "helix 6"  
 FT 162..183 /note= "helix 7"  
 FT Region 184..205 /note= "helix 8"  
 FT 206..227  
 FT Region

FT /note= "helix 9"  
 FT Region 228..249 /note= "helix 10"  
 FT 250..267 /note= "helix 11"  
 FT Region 268..289 /note= "helix 12"  
 FT 290..311 /note= "helix 13"  
 FT Region 312..333 /note= "helix 14"  
 FT 334..377 /note= "helix 15"  
 FT Region  
 PN WO9315198-A1.  
 XX  
 PD 05-AUG-1993.  
 XX  
 PF 26-JAN-1993; 93WO-FR00073.  
 XX  
 PR 27-JAN-1992; 92FR-0000806.  
 XX  
 PA (RHON ) RHONE POULENC RORER SA.  
 XX  
 PI Deneffe P, Guinet F, Latra M, Murry-Brelier A;  
 XX WPI; 1993-258676/32.  
 XX  
 PT New apolipoprotein AIV mutant polypeptide(s) - useful for drug  
 PT design and/or treatment of hypercholesterolaemia  
 XX  
 PS Claim 7; Page 31-32 and Page 5; 42pp; French.  
 XX  
 CC Human apolipoprotein AIV is a major component of chylomicrons in  
 CC lymph but is mainly found in the plasma in unassociated form. The  
 CC apo AIV mediates inverse transport of cholesterol. The invention  
 CC covers polypeptides which are derived from apoAIV by deletion of  
 CC at least 10 terminal amino acids, by deletion of a helix or pair  
 CC of helices, by addition of a heterologous polypeptide portion or  
 CC by a point mutation. Such mutants are useful in the design of  
 CC hypocholesterolaemic drugs to treat hypercholesterolaemia and  
 CC atherosclerosis. AAR39502 is a specifically claimed mutein and the  
 CC sequence has been compiled from the wild-type sequence (see AAR39443)  
 CC and the description given in the disclosure; the mutant sequence is  
 CC not shown in the specification.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 XX  
 SQ Sequence 377 AA;  
 Query Match 100.0%; Score 56; DB 14; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KALVOQMEQLRQ 12  
 DB 297 KALVOQMEQLRQ 308  
 XX  
 RESULT 23  
 AAR39501  
 ID AAR39501 standard; Protein; 377 AA.  
 XX  
 AC AAR39501;  
 DT 25-MAR-2003 (updated)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apoAIV mutein P(D5S).  
 XX  
 KM apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KM coronary disease; chylomicron; cholesterol transport;  
 KM plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.  
 XX


OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Protein 2..377 /note= "mature P(D5S)"  
FT Misc-difference 6 /note= "wild-type Asp5 is subst. by Ser"  
FT Region 14..40 /note= "helix 1"  
FT /note= "helix 1"  
FT Region 41..62 /note= "helix 2"  
FT /note= "helix 2"  
FT Region 63..95 /note= "helix 3"  
FT /note= "helix 3"  
FT Region 96..117 /note= "helix 4"  
FT /note= "helix 4"  
FT Region 118..139 /note= "helix 5"  
FT /note= "helix 5"  
FT Region 140..161 /note= "helix 6"  
FT /note= "helix 6"  
FT Region 162..183 /note= "helix 7"  
FT /note= "helix 7"  
FT Region 184..205 /note= "helix 8"  
FT /note= "helix 8"  
FT Region 206..227 /note= "helix 9"  
FT /note= "helix 9"  
FT Region 228..249 /note= "helix 10"  
FT /note= "helix 10"  
FT Region 250..267 /note= "helix 11"  
FT /note= "helix 11"  
FT Region 268..289 /note= "helix 12"  
FT /note= "helix 12"  
FT Region 290..311 /note= "helix 13"  
FT /note= "helix 13"  
FT Region 312..333 /note= "helix 14"  
FT /note= "helix 14"  
FT Region 334..377 /note= "helix 15"  
FT /note= "helix 15"  
XX  
XX WO9315198-A1.  
XX  
XX 05-AUG-1993. *Stone*  
XX  
XX PD  
XX 26-JAN-1993; 93WO-FR00073.  
XX  
XX PR 27-JAN-1992; 92FR-0000806.  
XX  
XX PA (RHON ) RHONE POULENC RORER SA.  
XX  
XX PI Deneffe P, Guinet F, Latta M, Murry-Brelrier A;  
XX  
XX DR WPI, 1993-258676/32.  
XX  
XX PS  
XX  
XX Claim 7; Page 31-32 and Page 5; 42pp; French.  
XX  
XX Human apolipoprotein AIV is a major component of chylomicrons in  
XX lymph but is mainly found in the plasma in unassociated form. The  
XX apo AIV mediates inverse transport of cholesterol. The invention  
XX covers polypeptides which are derived from apoAIV by deletion of  
XX at least 10 terminal amino acids, by deletion of a helix or pair  
XX of helices, by addition of a heterologous polypeptide portion or  
XX by a point mutation. Such mutants are useful in the design of  
XX hypocholesterolaemic drugs to treat hypercholesterolaemia and  
XX atherosclerosis. AAR39501 is a specifically claimed mutant and the  
XX sequence has been compiled from the wild-type sequence (see AAR39443)  
XX and the description given in the disclosure; the mutant sequence is  
XX not shown in the specification.  
XX (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 377 AA;  
XX  
XX

Query Match 100.0%; Score 56; DB 14; Length 377;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KALVOQMEQLRQ 12  
DB 297 KALVOQMEQLRQ 308  
RESULT 24  
AAR45242  
ID AAR45242 standard; Protein; 377 AA.  
XX  
XX AAR45242;  
XX  
XX AC  
XX 25-MAR-2003 (updated)  
XX DT 28-JAN-1994 (first entry)  
XX  
XX DE Human apoAIV mutain P(K178Y).  
XX  
XX KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
XX coronary disease; chylomicron; cholesterol transport;  
XX plaque deposition; lecithin-cholesterol acyltransferase; LCAT.  
XX  
XX OS Synthetic.  
XX  
XX FH Key Location/Qualifiers  
FT Protein 2..377 /note= "mature P(K178Y)"  
FT Misc-difference 179 /note= "wild-type Lys178 is subst. by Tyr"  
FT /note= "helix 1"  
FT Region 14..40 /note= "helix 1"  
FT /note= "helix 1"  
FT Region 41..62 /note= "helix 2"  
FT /note= "helix 2"  
FT Region 63..95 /note= "helix 3"  
FT /note= "helix 3"  
FT Region 96..117 /note= "helix 4"  
FT /note= "helix 4"  
FT Region 118..139 /note= "helix 5"  
FT /note= "helix 5"  
FT Region 140..161 /note= "helix 6"  
FT /note= "helix 6"  
FT Region 162..183 /note= "helix 7"  
FT /note= "helix 7"  
FT Region 184..205 /note= "helix 8"  
FT /note= "helix 8"  
FT Region 206..227 /note= "helix 9"  
FT /note= "helix 9"  
FT Region 228..249 /note= "helix 10"  
FT /note= "helix 10"  
FT Region 250..267 /note= "helix 11"  
FT /note= "helix 11"  
FT Region 268..289 /note= "helix 12"  
FT /note= "helix 12"  
FT Region 290..311 /note= "helix 13"  
FT /note= "helix 13"  
FT Region 312..333 /note= "helix 14"  
FT /note= "helix 14"  
FT Region 334..377 /note= "helix 15"  
FT /note= "helix 15"  
XX  
XX WO9315198-A1. *Stone*  
XX  
XX PD  
XX 05-AUG-1993.  
XX  
XX 26-JAN-1993; 93WO-FR00073.  
XX  
XX PR 27-JAN-1992; 92FR-0000806.  
XX  
XX PA (RHON ) RHONE POULENC RORER SA.  
XX  
XX PI Deneffe P, Guinet F, Latta M, Murry-Brelrier A;  
XX  
XX

XX WPI, 1993-258676/32.  
 DR New apo1ipoprotein AIV mutant polypeptide(s) - useful for drug  
 XX design and/or treatment of hypercholesterolaemia  
 PT  
 XX  
 PS Claim 7, Page 31-32 and Page 5, 42pp; French.  
 CC Human apolipoprotein AIV is a major component of chylomicrons in  
 CC lymph but is mainly found in the plasma in unassociated form. The  
 CC apo AIV mediates inverse transport of cholesterol. The invention  
 CC covers polypeptides which are derived from apoAIV by deletion of  
 CC at least 10 terminal amino acids, by deletion of a helix or pair  
 CC of helices, by addition of a heterologous polypeptide portion or  
 CC by a point mutation. Such mutants are useful in the design of  
 CC hypocholesterolaemic drugs to treat hypercholesterolaemia and  
 CC atherosclerosis. AAR45242 is a specifically claimed mutain and the  
 CC sequence has been compiled from the wild-type sequence (see AAR39443)  
 CC and the description given in the disclosure; the mutant sequence is  
 CC not shown in the specification.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 XX  
 SQ Sequence 377 AA;  
 Query Match 100.0%; Score 56; DB 14; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KALVOQMEOLRQ 12  
 DB 297 KALVOQMEOLRQ 308  
 RESULT 25  
 AAR45243  
 ID AAR45243 standard; Protein; 377 AA.  
 XX  
 AC AAR45243;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apoAIV mutain P(K178A).  
 XX  
 KM apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KM coronary disease; chylomicron; cholesterol transport;  
 KM plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 XX  
 FT Protein 2..377 /note= "mature P(K178A)"  
 FT Misc-difference 179 /note= "wild-type Lys178 is subseq. by Ala"  
 FT Region 14..40 /note= "helix 1"  
 FT Region 41..62 /note= "helix 2"  
 FT Region 63..95 /note= "helix 3"  
 FT Region 96..117 /note= "helix 4"  
 FT Region 118..139 /note= "helix 5"  
 FT Region 140..161 /note= "helix 6"  
 FT Region 162..183 /note= "helix 7"  
 FT Region 184..205 /note= "helix 8"  
 FT Region 206..227 /note= "helix 9"  
 FT

FT Region 228..249  
 FT /note= "helix 10"  
 FT Region 250..267  
 FT /note= "helix 11"  
 FT Region 268..289  
 FT /note= "helix 12"  
 FT Region 290..311  
 FT /note= "helix 13"  
 FT Region 312..333  
 FT /note= "helix 14"  
 FT Region 334..377  
 FT /note= "helix 15"  
 XX  
 PN WO9315198-A1.  
 XX  
 PD 05-AUG-1993.  
 XX  
 PF 26-JAN-1993; 93WO-FR00073.  
 XX  
 PR 27-JAN-1992; 92FR-0000806.  
 XX  
 PA (RHON ) RHONE POULENC RORER SA.  
 XX  
 PI Denefle P, Guinet F, Latra M, Murry-Brelier A;  
 XX  
 DR WPI, 1993-258676/32.  
 XX  
 PT New apo1ipoprotein AIV mutant polypeptide(s) - useful for drug  
 PT design and/or treatment of hypercholesterolaemia  
 XX  
 PS Claim 7, Page 31-32 and Page 5, 42pp; French.  
 XX  
 CC Human apolipoprotein AIV is a major component of chylomicrons in  
 CC lymph but is mainly found in the plasma in unassociated form. The  
 CC apo AIV mediates inverse transport of cholesterol. The invention  
 CC covers polypeptides which are derived from apoAIV by deletion of  
 CC at least 10 terminal amino acids, by deletion of a helix or pair  
 CC of helices, by addition of a heterologous polypeptide portion or  
 CC by a point mutation. Such mutants are useful in the design of  
 CC hypocholesterolaemic drugs to treat hypercholesterolaemia and  
 CC atherosclerosis. AAR45243 is a specifically claimed mutain and the  
 CC sequence has been compiled from the wild-type sequence (see AAR39443)  
 CC and the description given in the disclosure; the mutant sequence is  
 CC not shown in the specification.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 XX  
 SQ Sequence 377 AA;  
 Query Match 100.0%; Score 56; DB 14; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KALVOQMEOLRQ 12  
 DB 297 KALVOQMEOLRQ 308  
 RESULT 26  
 AAR45244  
 ID AAR45244 standard; Protein; 377 AA.  
 XX  
 AC AAR45244;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apoAIV mutain P(E230K).  
 XX  
 KM apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KM coronary disease; chylomicron; cholesterol transport;  
 KM plaque deposition; lecithin-cholesterol-acyltransferase; LCAT.  
 XX  
 OS Synthetic.



XX	Key	Location/Qualifiers
FH	Protein	2..377
FT		/note= "mature P(E230K)"
FT	Misc-difference	231
FT	Region	/note= "wild-type Glu230 is substcd. by Lys"
FT		14...40
FT	Region	/note= "helix 1"
FT		41..62
FT	Region	/note= "helix 2"
FT		63..95
FT	Region	/note= "helix 3"
FT		96..117
FT	Region	/note= "helix 4"
FT		118..139
FT	Region	/note= "helix 5"
FT		140..161
FT	Region	/note= "helix 6"
FT		162..183
FT	Region	/note= "helix 7"
FT		184..205
FT	Region	/note= "helix 8"
FT		206..227
FT	Region	/note= "helix 9"
FT		228..249
FT	Region	* /note= "helix 10"
FT		250..267
FT	Region	/note= "helix 11"
FT		268..289
FT	Region	/note= "helix 12"
FT		290..311
FT	Region	/note= "helix 13"
FT		312..333
FT	Region	/note= "helix 14"
FT		334..377
FT	Region	/note= "helix 15"
XX		
PM	W09315198-A1.	
PD	05-AUG-1993.	
PF	26-JAN-1993;	93WO-FR00073.
PR	27-JAN-1992;	92FR-0000806.
XX		
XX	(RHON ) RHONE POULENC RORER SA.	
PI	Deneffe P, Guinet F, Latta M, Murry-Brelier A;	
DR	WPI; 1993-258676/32.	
XX		
PT	New apo:Bipoprotein AIV mutant polypeptide(s) - useful for drug	
PT	design and/or treatment of hypercholesterolaemia	
XX		
XS	Claim 7; Page 31-32 and Page 5; 42pp; French.	
XX		
CC	Human apolipoprotein AIV is a major component of chylomicrons in lymph but is mainly found in the plasma in unassociated form. The apo AIV mediates inverse transport of cholesterol. The invention covers polypeptides which are derived from apoAIV by deletion of at least 10 terminal amino acids, by deletion of a helix or pair of helices, by addition of a heterologous polypeptide portion or by a point mutation. Such mutants are useful in the design of hypocholesterolemic drugs to treat hypercholesterolaemia and atherosclerosis. AAR45243 is a specifically claimed mutcin and the sequence has been compiled from the wild-type sequence (see AAR39443) and the description given in the disclosure; the mutant sequence is not shown in the specification. (Updated on 25-MAR-2003 to correct PN field.)	
CC		
CC		
XX	Sequence	377 AA;

Query Match      100.0%; Score 56; DB 14; Length 377;

Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY 1 KALVOOMEOLRQ 12  
|||  
Db 297 KALVOOMEOLRQ 308

RESULT 27  
AAB90664  
ID AAB90664 standard; Protein; 396 AA.  
XX AC  
AC AAB90664;  
XX DT  
DT 01-JUN-2001 (first entry)  
DE Human secreted protein, SEQ ID NO: 207.  
XX XX  
XX Human; secreted protein; immunomodulatory; antisclerotic;  
KW demarological; antiinflammatory; anti-HIV; cytostatic; cardiant;  
KW vascular; anti-angiogenic; ophthalmological; neuroprotectant;  
KW nocotropic; anticoagulant; antialzheimers; antiparkinsonian;  
KW antimicrobial; vulnery; vaccine; gene therapy; cancer;  
XX protein coordinate data; infection.

OS Homo sapiens.  
XX  
XX WO200121658-A1  
XX PD 29-MAR-2001.  
XX PF 22-SEP-2000; 2000WO-US26013.  
XX PR 24-SEP-1999; 99US-0155709.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Ni J., Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;  
PI Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;  
PI Young PE, Wei P, Florence KA;  
XX WP1; 2001-235311/24.

Nucleic acids encoding 32 human secreted polypeptides, useful for preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease and diabetic retinopathy -

Disclosure; Page 870-871; 890pp; English.

The present sequence is provided in a specification relating to nucleic acid molecules encoding 32 novel human secreted polypeptides. The nucleic acid molecules and polypeptides may be used in the prevention, diagnosis and treatment of diseases such as immune disorders (e.g. multiple sclerosis), systemic lupus erythematosus and human immuno-deficiency virus (HIV) infections), hyperproliferative disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases (e.g. Schmitzer syndrome, Chagas's cardiomyopathy and coronary arteriosclerosis), angiogenic disorders (e.g. corneal graft neovascularisation and diabetic retinopathy), neurological disorders (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease), infectious diseases and/or for promoting wound healing, regeneration and/or chemotaxis. The nucleic acid molecules may be used to produce the secreted polypeptides. They may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples. The polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators of their expression and activity.

Sequence 396 AA;

Query Match 100.0%; Score 56; DB 22; Length 396;  
Best Local Similarity 100.0%; Pred. NO. 0.11;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVQOMEQLRQ 12  
 |||||  
 DB 316 KALVQOMEQLRQ 327

## RESULT 28

AAW70729 standard; peptide; 28 AA.

AC AAR82729;

DT 03-MAY-1996 (first entry)

DE Human apolipoprotein A-IV fragment.

KM apolipoprotein; apo A-IV; appetite suppressant; food intake.

OS Synthetic.

PN W09525749-A2.

PD 28-SEP-1995.

PF 22-MAR-1995; 95WO-US03660.

PR 22-MAR-1994; 94US-0216537.

PA (RESE ) RESEARCH CORP TECHNOLOGIES INC.

PI Tso P;

DR WPI; 1995-344590/44.

PT Method to suppress appetite or inhibit food intake - by admin of the  
 PT 35 N-terminal amino acids of mature mammalian apo A-IV protein, or  
 PT analogues, homologues or fragments

PS Claim 8; Page 11; 100pp; English.

CC Novel eating suppressant peptides are provided which are derived  
 CC from apolipoprotein A-IV and have been made by solid phase peptide  
 CC synthesis. The peptides comprise at least a fragment of a 14 amino  
 CC acid sequence derived from the amino terminal portion of mature apo  
 CC A-IV. Smaller fragments (e.g. 3-13 amino acids) and larger peptides  
 CC (e.g. 15-30 amino acids) can also be used, as can homologues of  
 CC these sequences. Because of their small size, the peptides can pass  
 CC the blood brain barrier if necessary. They are not immunogenic, and  
 CC they provide a specific satiation signal.  
 CC The present sequence, which is corresponds substantially to residues  
 CC 316-346 of human apo A-IV precursor, is a specifically claimed example  
 CC of the new peptides.

SQ Sequence 28 AA;

Query Match 91.1%; Score 51; DB 16; Length 28;

Best Local Similarity 100.0%; Pred. No. 0.057; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALVQOMEQLRQ 12  
 |||||  
 DB 1 ALVQOMEQLRQ 11

## RESULT 29

AAW70713 standard; peptide; 28 AA.

AC AAW70713;

DT 28-JAN-1999 (first entry)

DE Apolipoprotein A-IV derived peptide analogue for suppressing appetite.

KW Analogue; human; rat; apolipoprotein A-IV; appetite suppression;  
 KW food intake inhibition.

OS Synthetic.

OS Homo sapiens.

PN US5840688-A.

PD 24-NOV-1998.

PF 22-MAR-1995; 95US-0408858.

PR 22-MAR-1995; 95US-0408858.

PA (RESE ) RESEARCH CORP TECHNOLOGIES INC.

PI Tso P;

DR WPI; 1999-034075/03.

PT Apolipoprotein A-IV peptides - useful as appetite suppressants  
 PS Claim 6; Column 5; 38pp; English.

CC AAW70704-90 represent peptide analogues of human and rat  
 CC apolipoprotein A-IV sequences. Most of the peptide analogues  
 CC correspond to a fundamental repeat unit (see W70791) in  
 CC the amino terminal region of apolipoprotein A-IV protein.  
 CC The peptide analogues suppress appetite and inhibit food intake.

SQ Sequence 28 AA;

Query Match 91.1%; Score 51; DB 20; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 0.057; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALVQOMEQLRQ 12  
 |||||  
 DB 1 ALVQOMEQLRQ 11

*Full*

## RESULT 30

AAO15885 standard; Protein; 396 AA.

AC AAO15885;

DT 30-JAN-2003 (first entry)

DE Human apolipoprotein A-IV (APOA4).

KW Human; gene therapy; single nucleotide polymorphism; SNP;

KW cytochrome C oxidase subunit VIb; COX6B; high serum cholesterol; GPI-1;

KW N-acetylglucosaminyl transferase component; cardiovascular disease; HDL;

OS Homo sapiens.

PN W0200272604-A2.

PD 19-SEP-2002.

PF 05-MAR-2002; 2002WO-US06728.

PR 09-MAR-2001; 2001US-0802640.

PA (SEQU-) SEQUENOM INC.

PI Braun A, Bansal A, Kleyen PW;

DR WPI; 2002-750478/81.  
 N-PSDB; ABT13007.

XX Detecting the presence or absence of an allelic variant of a  
 PT polymorphic region of COX6B and/or GPI-1 gene, useful for detecting a  
 PT predisposition to high serum cholesterol, low serum HDL and  
 PT cardiovascular disease -  
 XX  
 PS Disclosure; Page 117; 1999P; English.  
 CC The invention comprises methods of detecting the presence or absence of  
 CC at least one allelic variant of a polymorphic region of a gene associated  
 CC with cardiovascular disease. The invention specifically relates to  
 CC detecting the region of a cytochrome C oxidase subunit VIb (COX6B) gene  
 CC that is associated with high serum cholesterol, or the region of the  
 CC N-acetylglucosaminyl transferase component glycosylphosphatidylinositol-1  
 CC (GPI-1) gene that is associated with low serum high density lipoprotein  
 CC (HDL). The methods of the invention are useful for detecting a  
 CC predisposition to high serum cholesterol, low serum HDL and  
 CC cardiovascular disease. The methods are also useful for elucidating  
 CC pathological pathways, developing diagnostic assays and new drug  
 CC therapies for such disorders. The present amino acid sequence represents  
 CC a human protein that is associated with high serum cholesterol, low serum  
 CC HDL and/or cardiovascular disease.  
 CC  
 SQ Sequence 396 AA;  
 Query Match 91.1%; Score 51; DB 23; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.84; 0; Mismatches 0; Gaps 0;  
 Matches 11; Conservative 0; Indels 0; Gaps 0;  
 QY 1 KALVQOMEQLR 11  
 DB 316 KALVQOMEQLR 326  
 RESULT 31  
 AAU10860 ID AAU10860 standard; Protein; 396 AA.  
 AC AAU10860;  
 XX  
 DT 14-FEB-2002 (first entry)  
 DE Human apolipoprotein A-IV, APOA4.  
 KW Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;  
 KW haplotype; chromosome 11q23-qter; coronary heart disease; obesity;  
 KW atherosclerosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200177124-A2.  
 PD 18-OCT-2001.  
 PF 03-APR-2001; 2001WO-US10670.  
 PR 05-APR-2000; 2000US-194362P.  
 PA (GENA-) GENNAISSANCE PHARM INC.  
 PI Bentivegna SC, Choi JY, Klieem SE, Koshy B;  
 PI MPI; 2002-041281/05.  
 DR N-P8DB; AAS16512, AAS16513.  
 XX  
 PT New haplotypes of the human apolipoprotein A-IV gene, useful to  
 PT diagnose and treat disorders associated with its abnormal expression or  
 PT function such as coronary artery disease -  
 XX  
 PS Claim 29; Fig 3; 71pp; English.  
 CC The invention relates to haplotyping the human apolipoprotein A-IV  
 CC (APOA4) gene of an individual, comprising determining if the individual

CC has one of the APOA4 haplotypes or haplotype pairs fully defined in the  
 CC specification. Also disclosed are genotyping oligonucleotides (or  
 CC allele specific oligonucleotides, ASO) as well as methods for correlating  
 CC a particular haplotype pair with a trait e.g. obesity, in a population.  
 CC The APOA4 gene is located on chromosome 11q23-qter. The methods of the  
 CC invention are useful to diagnose and develop treatment for disorders  
 CC associated with abnormal APOA4 expression or function, for example  
 CC coronary heart disease and atherosclerosis. The APOA4 isoforms and  
 CC screened compounds are useful for the treatment of disorders associated  
 CC with abnormal APOA4 expression or function such as coronary artery  
 CC disease. The present sequence is the APOA4 protein.  
 CC  
 SQ Sequence 396 AA;  
 Query Match 91.1%; Score 51; DB 23; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.84; 0; Mismatches 0; Gaps 0;  
 Matches 11; Conservative 0; Indels 0; Gaps 0;  
 QY 1 KALVQOMEQLR 11  
 DB 316 KALVQOMEQLR 326  
 RESULT 32  
 AAU10861 ID AAU10861 standard; Protein; 396 AA.  
 AC AAU10861;  
 XX  
 DT 14-FEB-2002 (first entry)  
 DE Human apolipoprotein A-IV, APOA4, isoform #1.  
 KW Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;  
 KW haplotype; chromosome 11q23-qter; coronary heart disease; obesity;  
 KW atherosclerosis.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 13  
 FT note="Wild-type Val substituted by Met"  
 PN WO200177124-A2.  
 PD 18-OCT-2001.  
 PF 03-APR-2001; 2001WO-US10670.  
 PR 05-APR-2000; 2000US-194362P.  
 PA (GENA-) GENNAISSANCE PHARM INC.  
 PI Bentivegna SC, Choi JY, Klieem SE, Koshy B;  
 PI MPI; 2002-041281/05.  
 DR  
 XX  
 PT New haplotypes of the human apolipoprotein A-IV gene, useful to  
 PT diagnose and treat disorders associated with its abnormal expression or  
 PT function such as coronary artery disease -  
 XX  
 PS Claim 29; Page -; 71pp; English.  
 CC The invention relates to haplotyping the human apolipoprotein A-IV  
 CC (APOA4) gene of an individual, comprising determining if the individual  
 CC has one of the APOA4 haplotypes or haplotype pairs fully defined in the  
 CC specification. Also disclosed are genotyping oligonucleotides (or  
 CC allele specific oligonucleotides, ASO) as well as methods for correlating  
 CC a particular haplotype pair with a trait e.g. obesity, in a population.  
 CC The APOA4 gene is located on chromosome 11q23-qter. The methods of the  
 CC invention are useful to diagnose and develop treatment for disorders  
 CC associated with abnormal APOA4 expression or function, for example  
 CC coronary heart disease and atherosclerosis. The APOA4 isoforms and

CC screened compounds are useful for the treatment of disorders associated  
 CC with abnormal APOA4 expression or function such as coronary artery  
 CC disease. The present sequence is an isoform of the APOA4 protein.  
 CC Note: The present sequence is not shown in the specification but was  
 CC created by the indexer from the APOA4 sequence shown in figure 3.

XX SQ Sequence 396 AA;

Query Match 91.1%; Score 51; DB 23; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.84;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KALVOQMEQLR 11  
 |||||  
 DB 316 KALVOQMEQLR 326

RESULT 33

AAU10862 standard; Protein; 396 AA.

XX AAU10862;

XX 14-FEB-2002 (first entry)

XX Human apolipoprotein A-IV, APOA4, isoform #2.

XX Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;  
 KW haplotype; chromosome 11q23-qter; coronary heart disease; obesity;  
 KW atherosclerosis.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 111 /note= "Wild-type Ala substituted by Val"

XX WO200177124-A2.

XX 18-OCT-2001.

XX 03-APR-2001; 2001WO-US10670.

XX 05-APR-2000; 2000US-194362P.

XX (GENA-) GENAISSANCE PHARM INC.

XX Benticvegna SC, Choi JY, Klieem SE, Koshy B;

XX WPI; 2002-041281/05.

XX New haplotypes of the human apolipoprotein A-IV gene, useful to  
 PT diagnose and treat disorders associated with its abnormal expression or  
 PT function such as coronary artery disease -

XX Claim 29; Page - : 71pp; English.

XX The invention relates to haplotyping the human apolipoprotein A-IV  
 CC (APOA4) gene of an individual, comprising determining if the individual  
 CC has one of the APOA4 haplotypes or haplotype pairs fully defined in the  
 CC specification. Also disclosed are genotyping oligonucleotides (or  
 CC allele specific oligonucleotides, ASO) as well as methods for correlating  
 CC a particular haplotype pair with a trait e.g. obesity, in a population.  
 CC The APOA4 gene is located on chromosome 11q23-qter. The methods of the  
 CC invention are useful to diagnose and develop treatment for disorders  
 CC associated with abnormal APOA4 expression or function, for example  
 CC coronary heart disease and atherosclerosis. The APOA4 isoforms and  
 CC screened compounds are useful for the treatment of disorders associated  
 CC with abnormal APOA4 expression or function such as coronary artery  
 CC disease. The present sequence is an isoform of the APOA4 protein.  
 CC Note: The present sequence is not shown in the specification but was  
 CC created by the indexer from the APOA4 sequence shown in figure 3.

SQ Sequence 396 AA;

Query Match 91.1%; Score 51; DB 23; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.84;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KALVOQMEQLR 11  
 |||||  
 DB 316 KALVOQMEQLR 326

RESULT 34

AAU10863 standard; Protein; 396 AA.

XX AAU10863;

XX 14-FEB-2002 (first entry)

XX Human apolipoprotein A-IV, APOA4, isoform #3.

XX Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;  
 KW haplotype; chromosome 11q23-qter; coronary heart disease; obesity;  
 KW atherosclerosis.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 127 /note= "Wild-type Asn substituted by Lys"

XX WO200177124-A2.

XX 18-OCT-2001.

XX 03-APR-2001; 2001WO-US10670.

XX 05-APR-2000; 2000US-194362P.

XX (GENA-) GENAISSANCE PHARM INC.

XX Benticvegna SC, Choi JY, Klieem SE, Koshy B;

XX WPI; 2002-041281/05.

XX New haplotypes of the human apolipoprotein A-IV gene, useful to  
 PT diagnose and treat disorders associated with its abnormal expression or  
 PT function such as coronary artery disease -

XX Claim 29; Page - : 71pp; English.

XX The invention relates to haplotyping the human apolipoprotein A-IV  
 CC (APOA4) gene of an individual, comprising determining if the individual  
 CC has one of the APOA4 haplotypes or haplotype pairs fully defined in the  
 CC specification. Also disclosed are genotyping oligonucleotides (or  
 CC allele specific oligonucleotides, ASO) as well as methods for correlating  
 CC a particular haplotype pair with a trait e.g. obesity, in a population.  
 CC The APOA4 gene is located on chromosome 11q23-qter. The methods of the  
 CC invention are useful to diagnose and develop treatment for disorders  
 CC associated with abnormal APOA4 expression or function, for example  
 CC coronary heart disease and atherosclerosis. The APOA4 isoforms and  
 CC screened compounds are useful for the treatment of disorders associated  
 CC with abnormal APOA4 expression or function such as coronary artery  
 CC disease. The present sequence is an isoform of the APOA4 protein.  
 CC Note: The present sequence is not shown in the specification but was  
 CC created by the indexer from the APOA4 sequence shown in figure 3.

XX SQ Sequence 396 AA;

Query Match 91.1%; Score 51; DB 23; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.84;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVOQMEQLR 11  
 DB 316 KALVOQMEQLR 326

## RESULT 35

AAU10864 standard; Protein; 396 AA.

AAU10864;

14-FEB-2002 (first entry)

Human apolipoprotein A-IV, APOA4, isoform #4.

Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant; haplotype; chromosome 11q23-qter; coronary heart disease; obesity; atherosclerosis.

Homo sapiens.

Key Location/Qualifiers  
 FT Misc-difference 147 /note= "Wild-type Asn substituted by Ser"

MO200177124-A2.

18-OCT-2001.

03-APR-2001; 2001WO-US10670.

05-APR-2000; 2000US-194362P.

(GENA-) GENAISSANCE PHARM INC.

Bentivegna SC, Choi JY, Klieem SE, Koshy B;

WPI; 2002-041281/05.

New haplotypes of the human apolipoprotein A-IV gene, useful to diagnose and treat disorders associated with its abnormal expression or function such as coronary artery disease -  
 Claim 29; Page - ; 71pp; English.

The invention relates to haplotyping the human apolipoprotein A-IV (APOA4) gene of an individual, comprising determining if the individual has one of the APOA4 haplotypes or haplotype pairs fully defined in the specification. Also disclosed are genotyping oligonucleotides (or allele specific oligonucleotides, ASO) as well as methods for correlating a particular haplotype pair with a trait e.g. obesity, in a population. The APOA4 gene is located on chromosome 11q23-qter. The methods of the invention are useful to diagnose and develop treatment for disorders associated with abnormal APOA4 expression or function, for example coronary heart disease and atherosclerosis. The APOA4 isogenes and screened compounds are useful for the treatment of disorders associated with abnormal APOA4 expression or function such as coronary artery disease. The present sequence is an isoform of the APOA4 protein. Note: The present sequence is not shown in the specification but was created by the indexer from the APOA4 sequence shown in figure 3.

Sequence 396 AA;

Query Match 91.1%; Score 51; DB 23; Length 396;

Best Local Similarity 100.0%; Pred. No. 0.84; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVOQMEQLR 11  
 DB 316 KALVOQMEQLR 326

## RESULT 36

AAU10865 standard; Protein; 396 AA.

AAU10865;

14-FEB-2002 (first entry)

Human apolipoprotein A-IV, APOA4, isoform #5.

Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant; haplotype; chromosome 11q23-qter; coronary heart disease; obesity; atherosclerosis.

Homo sapiens.

Key Location/Qualifiers  
 FT Misc-difference 161 /note= "Wild-type Ala substituted by Ser"

MO200177124-A2.

18-OCT-2001.

03-APR-2001; 2001WO-US10670.

05-APR-2000; 2000US-194362P.

(GENA-) GENAISSANCE PHARM INC.

Bentivegna SC, Choi JY, Klieem SE, Koshy B;

WPI; 2002-041281/05.

New haplotypes of the human apolipoprotein A-IV gene, useful to diagnose and treat disorders associated with its abnormal expression or function such as coronary artery disease -  
 Claim 29; Page - ; 71pp; English.

The invention relates to haplotyping the human apolipoprotein A-IV (APOA4) gene of an individual, comprising determining if the individual has one of the APOA4 haplotypes or haplotype pairs fully defined in the specification. Also disclosed are genotyping oligonucleotides (or allele specific oligonucleotides, ASO) as well as methods for correlating a particular haplotype pair with a trait e.g. obesity, in a population. The APOA4 gene is located on chromosome 11q23-qter. The methods of the invention are useful to diagnose and develop treatment for disorders associated with abnormal APOA4 expression or function, for example coronary heart disease and atherosclerosis. The APOA4 isogenes and screened compounds are useful for the treatment of disorders associated with abnormal APOA4 expression or function such as coronary artery disease. The present sequence is an isoform of the APOA4 protein. Note: The present sequence is not shown in the specification but was created by the indexer from the APOA4 sequence shown in figure 3.

Sequence 396 AA;

Query Match 91.1%; Score 51; DB 23; Length 396;

Best Local Similarity 100.0%; Pred. No. 0.84; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVOQMEQLR 11  
 DB 316 KALVOQMEQLR 326

## RESULT 37

AAU10866 standard; Protein; 396 AA.

AAU10866;

14-FEB-2002 (first entry)

```

XX XX Human apolipoprotein A-IV, APOA4, isoform #6.
DE
XX
XX
KM Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;
KM haplotype; chromosome 11q23-qter; coronary heart disease; obesity;
KM atherosclerosis.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Misc-difference 193
FT /label= Met, Ile
FT /note= "replaces wild-type Val"
XX
XX WO200177124-A2.
XX
XX 18-OCT-2001.
XX
XX 03-APR-2001; 2001WO-US10670.
XX
XX 05-APR-2000; 2000US-194362P.
XX
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Bentivegna SC, Choi JY, Klieem SE, Koshy B;
XX
XX WPI; 2002-041281/05.
XX
XX New haplotypes of the human apolipoprotein A-IV gene, useful to
XX diagnose and treat disorders associated with its abnormal expression or
XX function such as coronary artery disease -
XX
XX Claim 29; Page - ; 71pp; English.
XX
XX The invention relates to haplotyping the human apolipoprotein A-IV
XX (APOA4) gene of an individual, comprising determining if the individual
XX has one of the APOA4 haplotypes or haplotype pairs fully defined in the
XX specification. Also disclosed are genotyping oligonucleotides (or
XX allele specific oligonucleotides, ASO) as well as methods for correlating
XX a particular haplotype pair with a trait e.g. obesity, in a population.
XX The APOA4 gene is located on chromosome 11q23-qter. The methods of the
XX invention are useful to diagnose and develop treatment for disorders
XX associated with abnormal APOA4 expression or function, for example
XX coronary heart disease and atherosclerosis. The APOA4 isogenes and
XX screened compounds are useful for the treatment of disorders associated
XX with abnormal APOA4 expression or function such as coronary artery
XX disease. The present sequence is an isoform of the APOA4 protein.
XX Note: The present sequence is not shown in the specification but was
XX created by the indexer from the APOA4 sequence shown in figure 3.
XX
XX Sequence 396 AA:
SQ
XX
XX Query Match 91.1%; Score 51; DB 23; Length 396;
XX Best Local Similarity 100.0%; Pred. No. 0.84;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 KALVOOMEQLR 11
XX |||||
XX 316 KALVOOMEQLR 326
XX
XX Db
XX
XX RESULT 38
XX AAU10867
XX ID AAU10867 strand; Protein; 396 AA.
XX
XX AAU10867;
XX
XX 14-FEB-2002 (first entry)
XX
XX Human apolipoprotein A-IV, APOA4, isoform #7.
XX
XX Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;
XX haplotype; chromosome 11q23-qter; coronary heart disease; obesity;
XX
XX
XX

```

```

KM atherosclerosis.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Misc-difference 353
FT /note= "wild-type Ser substituted by Ala"
XX
XX WO200177124-A2.
XX
XX 18-OCT-2001.
XX
XX 03-APR-2001; 2001WO-US10670.
XX
XX 05-APR-2000; 2000US-194362P.
XX
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Bentivegna SC, Choi JY, Klieem SE, Koshy B;
XX
XX WPI; 2002-041281/05.
XX
XX New haplotypes of the human apolipoprotein A-IV gene, useful to
XX diagnose and treat disorders associated with its abnormal expression or
XX function such as coronary artery disease -
XX
XX Claim 29; Page - ; 71pp; English.
XX
XX The invention relates to haplotyping the human apolipoprotein A-IV
XX (APOA4) gene of an individual, comprising determining if the individual
XX has one of the APOA4 haplotypes or haplotype pairs fully defined in the
XX specification. Also disclosed are genotyping oligonucleotides (or
XX allele specific oligonucleotides, ASO) as well as methods for correlating
XX a particular haplotype pair with a trait e.g. obesity, in a population.
XX The APOA4 gene is located on chromosome 11q23-qter. The methods of the
XX invention are useful to diagnose and develop treatment for disorders
XX associated with abnormal APOA4 expression or function, for example
XX coronary heart disease and atherosclerosis. The APOA4 isogenes and
XX screened compounds are useful for the treatment of disorders associated
XX with abnormal APOA4 expression or function such as coronary artery
XX disease. The present sequence is an isoform of the APOA4 protein.
XX Note: The present sequence is not shown in the specification but was
XX created by the indexer from the APOA4 sequence shown in figure 3.
XX
XX Sequence 396 AA:
SQ
XX
XX Query Match 91.1%; Score 51; DB 23; Length 396;
XX Best Local Similarity 100.0%; Pred. No. 0.84;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 KALVOOMEQLR 11
XX |||||
XX 316 KALVOOMEQLR 326
XX
XX Db
XX
XX RESULT 39
XX AAU10868
XX ID AAU10868 strand; Protein; 396 AA.
XX
XX AAU10868;
XX
XX 14-FEB-2002 (first entry)
XX
XX Human apolipoprotein A-IV, APOA4, isoform #8.
XX
XX Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;
XX haplotype; chromosome 11q23-qter; coronary heart disease; obesity;
XX atherosclerosis.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Misc-difference 367

```

```

FT      /note= "Wild-type Thr substituted by Ser"
XX
XX      WO200177124-A2.
XX
XX      18-OCT-2001.
XX
XX      03-APR-2001; 2001WO-US10670.
XX
XX      05-APR-2000; 2000US-194362P.
XX
XX      (GENA-) GENAISSANCE PHARM INC.
XX
XX      Bentivegna SC, Choi JY, Klem SE, Koshy B,
XX
XX      WPI; 2002-041281/05.
XX
XX      New haplotypes of the human apolipoprotein A-IV gene, useful to
XX      diagnose and treat disorders associated with its abnormal expression or
XX      function such as coronary artery disease
XX
XX      Claim 29; Page - ; 71pp; English.
XX
XX      The invention relates to haplotyping the human apolipoprotein A-IV
XX      (APOA4) gene of an individual, comprising determining if the individual
XX      has one of the APOA4 haplotypes or haplotype pairs fully defined in the
XX      specification. Also disclosed are genotyping oligonucleotides (or
XX      allele specific oligonucleotides, ASO) as well as methods for correlating
XX      a particular haplotype pair with a trait e.g. obesity, in a population.
XX      The APOA4 gene is located on chromosome 11q23-qter. The methods of the
XX      invention are useful to diagnose and develop treatment for disorders
XX      associated with abnormal APOA4 expression or function, for example
XX      coronary heart disease and atherosclerosis. The APOA4 isogenes and
XX      screened compounds are useful for the treatment of disorders associated
XX      with abnormal APOA4 expression or function such as coronary artery
XX      disease. The present sequence is an isoform of the APOA4 protein.
XX      Note: The present sequence is not shown in the specification but was
XX      created by the indexer from the APOA4 sequence shown in figure 3.
XX
XX      SQ      Sequence      396 AA;
XX
XX      Query Match      91.1%; Score 51; DB 23; Length 396;
XX      Best Local Similarity 100.0%; Pred. No. 0.84;
XX      Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      1 KALVOQMEQLR 11
XX      |||||
XX      316 KALVOQMEQLR 326
XX
XX      RESULT 40
XX      AAU10869
XX      ID      AAU10869 standard; Protein; 396 AA.
XX
XX      AC      AAU10869;
XX
XX      DT      14-FEB-2002 (first entry)
XX
XX      DE      Human apolipoprotein A-IV, APOA4, isoform #9.
XX
XX      KW      Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiac;
XX      haplotype; chromosome 11q23-qter; coronary heart disease; obesity;
XX      atherosclerosis.
XX
XX      OS      Homo sapiens.
XX
XX      Key      Location/Qualifiers
XX      FH      Misc-difference 380
XX      FT      /note= "Wild-type His substituted by Gln"
XX      PV      WO200177124-A2.
XX      PD      18-OCT-2001.
XX

```

```

PF      03-APR-2001; 2001WO-US10670.
XX
XX      05-APR-2000; 2000US-194362P.
XX
XX      (GENA-) GENAISSANCE PHARM INC.
XX
XX      Bentivegna SC, Choi JY, Klem SE, Koshy B,
XX
XX      WPI; 2002-041281/05.
XX
XX      New haplotypes of the human apolipoprotein A-IV gene, useful to
XX      diagnose and treat disorders associated with its abnormal expression or
XX      function such as coronary artery disease
XX
XX      Claim 29; Page - ; 71pp; English.
XX
XX      The invention relates to haplotyping the human apolipoprotein A-IV
XX      (APOA4) gene of an individual, comprising determining if the individual
XX      has one of the APOA4 haplotypes or haplotype pairs fully defined in the
XX      specification. Also disclosed are genotyping oligonucleotides (or
XX      allele specific oligonucleotides, ASO) as well as methods for correlating
XX      a particular haplotype pair with a trait e.g. obesity, in a population.
XX      The APOA4 gene is located on chromosome 11q23-qter. The methods of the
XX      invention are useful to diagnose and develop treatment for disorders
XX      associated with abnormal APOA4 expression or function, for example
XX      coronary heart disease and atherosclerosis. The APOA4 isogenes and
XX      screened compounds are useful for the treatment of disorders associated
XX      with abnormal APOA4 expression or function such as coronary artery
XX      disease. The present sequence is an isoform of the APOA4 protein.
XX      Note: The present sequence is not shown in the specification but was
XX      created by the indexer from the APOA4 sequence shown in figure 3.
XX
XX      SQ      Sequence      396 AA;
XX
XX      Query Match      91.1%; Score 51; DB 23; Length 396;
XX      Best Local Similarity 100.0%; Pred. No. 0.84;
XX      Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      1 KALVOQMEQLR 11
XX      |||||
XX      316 KALVOQMEQLR 326
XX
XX      Search completed: September 8, 2003, 16:04:59
XX      Job time : 41 secs
XX

```

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2003, 16:04:13 ; Search time 22 Seconds  
(without alignments)  
23.079 Million cell updates/sec

Title: US-09-991-809-1

Perfect score: 56  
Sequence: 1 KALVQOMQLRQ 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/CTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	91.1	28	2	US-08-408-858A-10
2	43	76.8	30	2	US-08-408-858A-7
3	38	67.9	500	4	US-09-252-991A-30255
4	36	64.3	434	4	US-09-328-352-4261
5	35	62.5	732	4	US-08-671-757A-7
6	35	62.5	732	4	US-08-671-757A-8
7	35	62.5	956	4	US-09-914-259-17
8	35	62.5	957	4	US-09-914-259-16
9	35	62.5	10182	4	US-09-134-001C-3159
10	34	60.7	724	4	US-09-328-352-7710
11	34	60.7	1085	1	US-08-431-080-28
12	34	60.7	1085	2	US-08-938-534-28
13	34	60.7	1085	4	US-09-345-294-28
14	33	58.9	230	4	US-09-252-991A-23778
15	33	58.9	478	4	US-09-252-991A-25711
16	33	58.9	678	4	US-09-252-991A-20693
17	33	58.9	1507	4	US-09-914-259-17
18	32	57.1	67	4	US-09-695-458-12
19	32	57.1	76	4	US-09-695-458-12
20	32	57.1	231	4	US-09-328-352-4493
21	32	57.1	263	4	US-09-328-352-4493
22	32	57.1	304	4	US-09-695-458-3
23	32	57.1	307	3	US-09-267-031-6
24	32	57.1	325	3	US-09-267-031-14
25	32	57.1	330	3	US-09-267-031-12
26	32	57.1	330	3	US-09-267-031-16
27	32	57.1	331	4	US-09-695-458-2

28	32	57.1	413	4	US-09-328-352-4779	Sequence 4779, Ap
29	32	57.1	438	1	US-08-111-939-17	Sequence 17, Appl
30	32	57.1	458	3	US-09-233-989-9	Sequence 9, Appl
31	32	57.1	468	4	US-09-107-532A-5304	Sequence 5304, Ap
32	32	57.1	876	1	US-08-785-071A-2	Sequence 2, Appl
33	32	57.1	876	3	US-09-012-872-2	Sequence 2, Appl
34	32	57.1	893	4	US-09-328-352-6626	Sequence 6626, Ap
35	32	57.1	955	1	US-08-006-676B-1	Sequence 1, Appl
36	32	57.1	955	2	US-08-282-845-2	Sequence 2, Appl
37	32	57.1	955	2	US-08-428-414A-3	Sequence 3, Appl
38	32	57.1	955	5	PCT-US94-00324-1	Sequence 1, Appl
39	32	57.1	983	4	US-09-134-001C-3814	Sequence 3814, Ap
40	31	55.4	54	2	US-08-319-866-3	Sequence 3, Appl
41	31	55.4	88	4	US-09-107-532A-6155	Sequence 6155, Ap
42	31	55.4	214	2	US-08-852-809-3	Sequence 3, Appl
43	31	55.4	218	4	US-09-328-352-7752	Sequence 7752, Ap
44	31	55.4	222	4	US-09-328-352-7790	Sequence 7790, Ap
45	31	55.4	227	4	US-09-328-352-6938	Sequence 6938, Ap

## ALIGNMENTS

RESULT 1  
US-08-408-858A-10  
; Sequence 10, Application US/08408858A  
; Patent No. 5840688  
; GENERAL INFORMATION:  
; APPLICANT: TBO, Patrick  
; TITLE OF INVENTION: EATING SUPPRESSANT PEPTIDES  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/408,858A  
; FILING DATE: 22-MAR-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digilio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 90212  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 516-742-4343  
; TELEFAX: 516-742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-408-858A-10

Query Match 91.1%; Score 51; DB 2; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KALVQOMQLRQ 12  
DB 1 KALVQOMQLRQ 11

*aligned*

RESULT 2  
US-08-408-858A-7  
Sequence 7, Application US/08408858A  
Patent No. 5840688

GENERAL INFORMATION:

APPLICANT: TSO, Patrick  
TITLE OF INVENTION: EATING SUPPRESSANT PEPTIDES  
NUMBER OF SEQUENCES: 88  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/408,858A  
FILING DATE: 22-MAR-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: DIGLIO, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 90212  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-408-858A-7

Query Match 76.8%; Score 43; DB 2; Length 30;  
Best Local Similarity 81.8%; Pred. No. 0.4;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALVQOMEQLRQ 12  
Db 1 ALVQOMEKFRQ 11

RESULT 3  
US-09-252-991A-30255  
Sequence 30255, Application US/09252991A  
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 30255  
LENGTH: 500  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30255

Query Match 67.9%; Score 38; DB 4; Length 500;

Best Local Similarity 80.0%; Pred. No. 50;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVQOMEQLRQ 12  
Db 365 LVQOMEQLRQ 374

RESULT 4  
US-09-328-352-4261  
Sequence 4261, Application US/09328352  
Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328,352

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 4261

LENGTH: 434

TYPE: PRT

ORGANISM: Acinetobacter baumannii

US-09-328-352-4261

Query Match 64.3%; Score 36; DB 4; Length 434;  
Best Local Similarity 58.3%; Pred. No. 95;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KALVQOMEQLRQ 12  
Db 51 KALVQOEORQVOQ 62

RESULT 5  
US-08-671-757A-7  
Sequence 7, Application US/08671757A  
Patent No. 6476213

GENERAL INFORMATION:

APPLICANT: Suerbaum, Sebastian  
TITLE OF INVENTION: Cloning and Characterization Production

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegun, Henderson, Farabow, Garrett &

Dunnet  
STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/08/671,757A

CURRENT FILING DATE: 16-Aug-2001

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 02356.0073-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 408-4400  
TELEFAX: (202) 408-4400  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 732 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-08-671-757A-7

Query Match 62.5%; Score 35; DB 4; Length 732;  
Best Local Similarity 72.7%; Pred. No. 2.4e+02;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 KALVQOMEQLR 11  
Db 695 KALSNOMEQAR 705

RESULT 6  
US-08-671-757A-8  
Sequence 8, Application US/08671757A  
Patent No. 6476213

GENERAL INFORMATION:

APPLICANT: Suerbaum, Sebastian

TITLE OF INVENTION: Cloning and Characterization Production  
of Aflagellate Strains

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:  
ADDRESS: Finnegan, Henderson, Farabow, Garrett &  
Dunner

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/671,757A

FILING DATE: 16-Aug-2001

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Kenneth J.

REGISTRATION NUMBER: 25,146

REFERENCE/DOCKET NUMBER: 02356, 0073-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 408-4000

TELEFAX: (202) 408-4400

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 732 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-08-671-757A-8

Query Match 62.5%; Score 35; DB 4; Length 732;  
Best Local Similarity 72.7%; Pred. No. 2.4e+02;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 KALVQOMEQLR 11  
Db 695 KALSNOMEQAR 705

RESULT 7  
US-09-914-259-17  
Sequence 17, Application US/09914259  
Patent No. 6495336

GENERAL INFORMATION:

APPLICANT: Makowski, Lee

APPLICANT: Hyman, Paul

APPLICANT: Williams, Mark  
TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES  
FILE REFERENCE: 8471-010-999  
CURRENT APPLICATION NUMBER: US/09/914,259  
CURRENT FILING DATE: 2000-11-21  
NUMBER OF SEQ ID NOS: 180  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 17  
LENGTH: 956  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-914-259-17

Query Match 62.5%; Score 35; DB 4; Length 956;  
Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KALVQOMEQLRQ 12  
Db 700 KALQOMESHRE 711

RESULT 8  
US-09-914-259-16  
Sequence 16, Application US/09914259  
Patent No. 6495336

GENERAL INFORMATION:

APPLICANT: Makowski, Lee

APPLICANT: Hyman, Paul

APPLICANT: Williams, Mark

TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES

FILE REFERENCE: 8471-010-999

CURRENT APPLICATION NUMBER: US/09/914,259

CURRENT FILING DATE: 2000-11-21

NUMBER OF SEQ ID NOS: 180

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 16

LENGTH: 957

TYPE: PRT

ORGANISM: Homo sapiens

US-09-914-259-16

Query Match 62.5%; Score 35; DB 4; Length 957;  
Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KALVQOMEQLRQ 12  
Db 701 KALQOMESHRE 712

RESULT 9  
US-09-134-001C-3159  
Sequence 3159, Application US/09134001C  
Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NOCTIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779

PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 3159

LENGTH: 10182

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

US-09-134-001C-3159

Query Match 62.5%; Score 35; DB 4; Length 10182;  
Best Local Similarity 58.3%; Pred. No. 3.5e+03;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KALVQMEQLRQ 12  
Db 6953 KALVQMEQLRQ 6964

RESULT 10

US-09-328-352-7710

Sequence 7710, Application US/09328352

Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328,352

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 7710

LENGTH: 724

TYPE: PRT

ORGANISM: Acinetobacter baumannii

US-09-328-352-7710

Query Match 60.7%; Score 34; DB 4; Length 724;  
Best Local Similarity 58.3%; Pred. No. 3.5e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KALVQMEQLRQ 12  
Db 107 KGLVQLRLRK 118

RESULT 11

US-08-431-080-28

Sequence 28, Application US/08431080

Patent No. 5698686

GENERAL INFORMATION:

APPLICANT: Gottschling, Daniel E.

APPLICANT: Singer, Miriam S.

TITLE OF INVENTION: Telomerase Compositions and Methods

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: TEXAS

COUNTRY: UNITED STATES OF AMERICA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/431,080

FILING DATE: Concurrently Herewith

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: SN 08/326,781

FILING DATE: October 20, 1994

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Parker, David L.

REGISTRATION NUMBER: 32,165

REFERENCE/DOCKET NUMBER: ARCD:155/PAR

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679

TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1085 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

Query Match 60.7%; Score 34; DB 1; Length 1085;  
Best Local Similarity 54.5%; Pred. No. 5.3e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALVQMEQLRQ 12  
Db 740 ALVQMEQLRQ 750

RESULT 12

US-08-938-534-28

Sequence 28, Application US/08938534

Patent No. 5916752

GENERAL INFORMATION:

APPLICANT: Gottschling, Daniel E.

APPLICANT: Singer, Miriam S.

TITLE OF INVENTION: Telomerase Compositions and Methods

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: TEXAS

COUNTRY: UNITED STATES OF AMERICA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC Compatible

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/938,534

FILING DATE: 26-SEP-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/431,080

FILING DATE:

APPLICATION NUMBER: SN 08/326,781

FILING DATE: October 20, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Parker, David L.

REGISTRATION NUMBER: 32,165

REFERENCE/DOCKET NUMBER: ARCD:155/PAR

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679

TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 1085 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-938-534-28

Query Match 60.7%; Score 34; DB 2; Length 1085;  
Best Local Similarity 54.5%; Pred. No. 5.3e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALVQMEQLRQ 12  
Db 740 ALVQMEQLRQ 750

RESULT 13

US-09-345-294-28  
; Sequence 28, Application US/09345294  
; Patent No. 6387619  
; GENERAL INFORMATION:  
; APPLICANT: Gottschling, Daniel E.  
; SINGER, Miriam S.  
; TITLE OF INVENTION: Telomerase Compositions and Methods  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: TEXAS  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/345,294  
; FILING DATE: 30-Jun-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/431,080  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parker, David L.  
; REGISTRATION NUMBER: 32,165  
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (512) 418-3000  
; TELEFAX: (713) 789-2679  
; TELEX: 79-0924  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1085 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
US-09-345-294-28  
  
Query Match 60.7%; Score 34; DB 4; Length 1085;  
Best Local Similarity 54.5%; Pred. No. 5.3e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 ALVQOMEQLRQ 12  
DB 740 ALIQEHQQLRE 750  
  
RESULT 14  
US-09-252-991A-23778  
; Sequence 23778, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23778  
; LENGTH: 230  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-23778  
Query Match 58.9%; Score 33; DB 4; Length 230;  
Best Local Similarity 50.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 KALVQOMEQLRQ 12  
DB 181 QALVQDLDDLKQ 192  
  
RESULT 15  
US-09-252-991A-25711  
; Sequence 25711, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 25711  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25711  
  
Query Match 58.9%; Score 33; DB 4; Length 478;  
Best Local Similarity 66.7%; Pred. No. 3.4e+02;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 KALVQOMEQLRQ 12  
DB 177 QALVQPEDDLQ 188  
  
RESULT 16  
US-09-252-991A-20693  
; Sequence 20693, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20693  
; LENGTH: 678  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20693  
  
Query Match 58.9%; Score 33; DB 4; Length 678;  
Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 KALVQOMEQLRQ 12  
DB 620 KGLAQETQOVRO 631

```
RESULT 17
US-09-914-259-37
; Sequence 37, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 1507
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-914-259-37

Query Match          58.9%; Score 32; DB 4; Length 1507;
Best Local Similarity 54.5%; Pred. No. 1,1e+03;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      2 ALVQQMELRQ 12
      |||:|||||
Db      884 AICQQLQQLQ 894

RESULT 18
US-09-695-458-12
; Sequence 12, Application US/09695458
; Patent No. 6380361
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Tate, David W.
; TITLE OF INVENTION: Educational Kit and Method Containing No. 6380361el Alpha Helical
; FILE REFERENCE: 99-78
; CURRENT APPLICATION NUMBER: US/09/695,458
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/162,623
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-695-458-12

Query Match          57.1%; Score 32; DB 4; Length 67;
Best Local Similarity 58.3%; Pred. No. 68;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 KALVQMEQLRQ 12
      |||:|||||
Db      38 EAAAOQHQLQ 49

RESULT 19
US-09-695-458-18
; Sequence 18, Application US/09695458
; Patent No. 6380361
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Tate, David W.
; TITLE OF INVENTION: Educational Kit and Method Containing No. 6380361el Alpha Helical
; FILE REFERENCE: 99-78
; CURRENT APPLICATION NUMBER: US/09/695,458
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/162,623
```

```
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-695-458-18

Query Match          57.1%; Score 32; DB 4; Length 76;
Best Local Similarity 63.3%; Pred. No. 77;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 KALVQMEQLRQ 12
      |||:|||||
Db      11 EAAAOQHQLQ 22

RESULT 20
US-09-328-352-4493
; Sequence 4493, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4493
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4493

Query Match          57.1%; Score 32; DB 4; Length 231;
Best Local Similarity 45.5%; Pred. No. 2,4e+02;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      1 KALVQMEQLR 11
      |||:|||||
Db      177 KPLQKIBQIK 187

RESULT 21
US-09-328-352-6598
; Sequence 6598, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6598
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6598

Query Match          57.1%; Score 32; DB 4; Length 263;
Best Local Similarity 63.6%; Pred. No. 2,7e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      2 ALVQQMELRQ 12
      |||:|||||
Db      87 ALRRDEQLRQ 97

RESULT 22
```

US-09-695-458-3  
; Sequence 3, Application US/09695458  
; Patent No. 6380361  
; GENERAL INFORMATION:  
; APPLICANT: Conklin, Darrell C.  
; TITLE OF INVENTION: Educational Kit and Method Containing No. 6380361el Alpha Helical  
; FILE REFERENCE: 99-78  
; CURRENT APPLICATION NUMBER: US/09/695,458  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: 60/162,623  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 304  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-695-458-3

Query Match 57.1%; Score 32; DB 4; Length 304;  
Best Local Similarity 58.3%; Pred. No. 3.2e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KALVQOMEQLRQ 12  
DB 232 EAAQOHEQLQO 243

RESULT 23  
US-09-267-031-6  
; Sequence 6, Application US/09267031  
; Patent No. 6137031  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Yuelin  
; APPLICANT: Kinkema, Mark  
; APPLICANT: Dong, Ximinian  
; APPLICANT: Ronald, Pamela  
; APPLICANT: Chern, Mawsheng  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: DNA Binding Proteins That Interact With NPRI  
; FILE REFERENCE: 023070-092500US  
; CURRENT APPLICATION NUMBER: US/09/267,031  
; CURRENT FILING DATE: 1999-03-11  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 307  
; TYPE: PRT  
; ORGANISM: Oryza sp.  
US-09-267-031-6

Query Match 57.1%; Score 32; DB 3; Length 307;  
Best Local Similarity 63.6%; Pred. No. 3.2e+02;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 KALVQOMEQLR 11  
DB 43 KAYVQOLENSR 53

RESULT 24  
US-09-267-031-14  
; Sequence 14, Application US/09267031  
; Patent No. 6137031  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Yuelin  
; APPLICANT: Kinkema, Mark  
; APPLICANT: Dong, Ximinian  
; APPLICANT: Ronald, Pamela  
; APPLICANT: Chern, Mawsheng

; APPLICANT: The Regents of the University of California  
; APPLICANT: Duke University  
; TITLE OF INVENTION: DNA Binding Proteins That Interact With NPRI  
; FILE REFERENCE: 023070-092500US  
; CURRENT APPLICATION NUMBER: US/09/267,031  
; CURRENT FILING DATE: 1999-03-11  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 325  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-267-031-14

Query Match 57.1%; Score 32; DB 3; Length 325;  
Best Local Similarity 63.6%; Pred. No. 3.4e+02;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KALVQOMEQLR 11  
DB 60 KAYVQOLENSR 70

RESULT 25  
US-09-267-031-12  
; Sequence 12, Application US/09267031  
; Patent No. 6137031  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Yuelin  
; APPLICANT: Kinkema, Mark  
; APPLICANT: Dong, Ximinian  
; APPLICANT: Ronald, Pamela  
; APPLICANT: Chern, Mawsheng  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: DNA Binding Proteins That Interact With NPRI  
; FILE REFERENCE: 023070-092500US  
; CURRENT APPLICATION NUMBER: US/09/267,031  
; CURRENT FILING DATE: 1999-03-11  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 330  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-267-031-12

Query Match 57.1%; Score 32; DB 3; Length 330;  
Best Local Similarity 63.6%; Pred. No. 3.4e+02;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KALVQOMEQLR 11  
DB 66 KAYVQOLENSR 76

RESULT 26  
US-09-267-031-16  
; Sequence 16, Application US/09267031  
; Patent No. 6137031  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Yuelin  
; APPLICANT: Kinkema, Mark  
; APPLICANT: Dong, Ximinian  
; APPLICANT: Ronald, Pamela  
; APPLICANT: Chern, Mawsheng  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: DNA Binding Proteins That Interact With NPRI  
; FILE REFERENCE: 023070-092500US  
; CURRENT APPLICATION NUMBER: US/09/267,031  
; CURRENT FILING DATE: 1999-03-11  
; NUMBER OF SEQ ID NOS: 26

SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 16  
LENGTH: 330  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-267-031-16

Query Match 57.1% Score 32; DB 3; Length 330;  
Best Local Similarity 63.6% Pred. No. 3.4e+02;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KALVQOMQLR 11  
Db 66 KAYVOOLENSR 76

RESULT 27  
US-09-695-458-2  
Sequence 2, Application US/09695458  
Patent No. 6380361  
GENERAL INFORMATION:  
APPLICANT: Tatf, David W.  
TITLE OF INVENTION: Educational Kit and Method Containing No. 6380361e1 Alpha Helical  
FILE REFERENCE: 99-78  
CURRENT APPLICATION NUMBER: US/09/695,458  
CURRENT FILING DATE: 2001-09-25  
PRIOR FILING DATE: 1999-10-29  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 331  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-695-458-2

Query Match 57.1% Score 32; DB 4; Length 331;  
Best Local Similarity 58.3% Pred. No. 3.4e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KALVQOMQLRQ 12  
Db 259 EAAAOQHEDLQ 270

RESULT 28  
US-09-328-352-4779  
Sequence 4779, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 4779  
LENGTH: 413  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-4779

Query Match 57.1% Score 32; DB 4; Length 413;  
Best Local Similarity 50.0% Pred. No. 4.3e+02;  
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVQOMQL 10  
Db 306 KSIIVQLQL 315

RESULT 29  
US-08-111-939-17  
Sequence 17, Application US/08111939  
Patent No. 5460951  
GENERAL INFORMATION:  
APPLICANT: Kawai, Shinji  
APPLICANT: Takeshita, Sunao  
APPLICANT: Okazaki, Makoto.  
APPLICANT: Amano, Egon  
TITLE OF INVENTION: Bone-Related Carboxypeptidase-Like  
TITLE OF INVENTION: Protein and Process for its Production  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/111,939  
FILING DATE: 26-AUG-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 324033/92  
FILING DATE: 03-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 230029/92  
FILING DATE: 28-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Forman, David S.  
REGISTRATION NUMBER: 33,694  
REFERENCE/DOCKET NUMBER: 02481.1321-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4000  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 438 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-111-939-17

Query Match 57.1% Score 32; DB 1; Length 438;  
Best Local Similarity 50.0% Pred. No. 4.6e+02;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KALVQOMQLRQ 12  
Db 309 EALIOFLQVHQ 320

RESULT 30  
US-09-233-989-9  
Sequence 9, Application US/09233989  
Patent No. 6248527  
GENERAL INFORMATION:  
APPLICANT: Chen, Hong  
APPLICANT: Meyer, Joanne  
TITLE OF INVENTION: Method of Detecting Risk of Type II Diabetes Based on  
FILE REFERENCE: Mutations Found in Carboxypeptidase E  
CURRENT APPLICATION NUMBER: US/09/233,989  
CURRENT FILING DATE: 1999-01-19  
EARLIER APPLICATION NUMBER: 60/105,102



EARLIER FILING DATE: 1998-10-21  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 9  
LENGTH: 458  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: carboxypeptidase homolog -- CPN  
US-09-233-989-9

Query Match 57.1%; Score 32; DB 3; Length 458;  
Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KALVQOMEQ 12  
Db 329 KALVQOMEQ 340

## RESULT 31

US-09-107-532A-5304  
Sequence 5304, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Denke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5304:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 468 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...468  
SEQUENCE DESCRIPTION: SEQ ID NO: 5304:  
US-09-107-532A-5304  
Query Match 57.1%; Score 32; DB 4; Length 468;  
Best Local Similarity 55.6%; Pred. No. 4.9e+02;

Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KALVQOMEQ 9  
Db 30 KSLIQOLEQ 38

## RESULT 32

US-08-785-071A-2  
Sequence 2, Application US/08785071A  
Patent No. 576750  
GENERAL INFORMATION:  
APPLICANT: Hodgson, John  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: No. 576750e1 crna Synthetase  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/785,071A  
FILING DATE: 17-JUN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9601099.6  
FILING DATE: 19-JAN-1996  
APPLICATION NUMBER: 9622617.0  
FILING DATE: 27-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31355-6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 876 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-785-071A-2  
Query Match 57.1%; Score 32; DB 1; Length 876;  
Best Local Similarity 66.7%; Pred. No. 9.3e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVQOMEQ 9  
Db 742 KALVQOMEQ 750

## RESULT 33

US-09-012-872-2  
Sequence 2, Application US/09012872  
Patent No. 6060294  
GENERAL INFORMATION:  
APPLICANT: Hodgson, John  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: No. 6060294e1 crna Synthetase  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:

```

; ADDRESS: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/012,872
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/785,071
; FILING DATE: 17-JAN-1997
; APPLICATION NUMBER: 9601099.6
; FILING DATE: 19-JAN-1996
; APPLICATION NUMBER: 9622617.0
; FILING DATE: 27-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P33355-6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 876 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-09-012-872-2

Query Match      57.1%; Score 32; DB 3; Length 876;
Best Local Similarity 66.7%; Pred. No. 9.3e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 KALVQMEQLRQ 9
      |||::|||
Db      742 KALKKQLRQ 750

RESULT 34
US-09-328-352-6626
; Sequence 6626, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6626
; LENGTH: 893
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
;
US-09-328-352-6626

Query Match      57.1%; Score 32; DB 4; Length 893;
Best Local Similarity 63.6%; Pred. No. 9.4e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 ALVQMEQLRQ 12
      |||::|||
Db      760 ALQKQTEQLNQ 770
```

```

RESULT 35
US-08-006-676B-1
; Sequence 1, Application US/08006676B
; Patent No. 5411665
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven
; TITLE OF INVENTION: Diagnosis of Leishmaniasis
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeffrey B. Oster
; STREET: 8339 SE 57th Street
; CITY: Mercer Island
; STATE: Washington
; COUNTRY: USA
; ZIP: 98040-4906
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/006,676B
; FILING DATE: 15-JAN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oster, Jeffrey B.
; REGISTRATION NUMBER: 32,585
; REFERENCE/DOCKET NUMBER: REED-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 232 7845
; TELEFAX: (206) 236 0205
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 955 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-006-676B-1

Query Match      57.1%; Score 32; DB 1; Length 955;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 KALVQMEQLRQ 12
      :|||::|||
Db      408 RELEQMEDMRQ 419

RESULT 36
US-08-282-845-2
; Sequence 2, Application US/08282845
; Patent No. 5719263
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: A 230Kd Antigen Present in Leishmania
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Apple Macintosh
; SOFTWARE: Microsoft Word for Macintosh 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,845
```

FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/006,676  
FILING DATE: JANUARY 15, 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,693  
REFERENCE/DOCKET NUMBER: 5004-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)233-0644  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 955 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-282-845-2

Query Match 57.1%; Score 32; DB 1; Length 955;  
Best Local Similarity 50.0%; Pred. No. 1e+03;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLRQ 12  
:|:|:|:|:  
Db 408 RELEQMEDMRQ 419

## RESULT 37

US-08-428-414A-3  
; Sequence 3, Application US/08428414A  
; Patent No. 5912166

GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
LEISHMANIASIS  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,414A  
FILING DATE: 21-APR-1995  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Kadlecik, Ann T.  
REGISTRATION NUMBER: 39,244  
REFERENCE/DOCKET NUMBER: 210121.407  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
TELEX: 3723836 SEDDANDBERRY  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 955 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-428-414A-3

Query Match 57.1%; Score 32; DB 2; Length 955;  
Best Local Similarity 50.0%; Pred. No. 1e+03;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 KALVOQMEQLRQ 12  
:|:|:|:|:  
Db 408 RELEQMEDMRQ 419

## RESULT 38

PCT-US94-00324-1  
; Sequence 1, Application PC/TUS9400324  
; GENERAL INFORMATION:

APPLICANT: Reed, Steven  
TITLE OF INVENTION: Diagnosis of Leishmaniasis  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98101

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Apple System 7.1  
SOFTWARE: Microsoft Word, version 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/00324  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/006,676  
FILING DATE: 15-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,693  
REFERENCE/DOCKET NUMBER: 5004-WO

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430  
TELEFAX: (206) 233-0644

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 955 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
PCT-US94-00324-1

Query Match 57.1%; Score 32; DB 5; Length 955;  
Best Local Similarity 50.0%; Pred. No. 1e+03;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLRQ 12  
:|:|:|:|:  
Db 408 RELEQMEDMRQ 419

RESULT 39  
US-09-134-001C-3814  
; Sequence 3814, Application US/09134001C  
; Patent No. 6380370

GENERAL INFORMATION:  
APPLICANT: Lynn Doucelte-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674

Query Match 57.1%; Score 32; DB 5; Length 955;  
Best Local Similarity 50.0%; Pred. No. 1e+03;

; SEQ ID NO 3814  
; LENGTH: 983  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3814

Query Match 57.1%; Score 32; DB 4; Length 983;  
Best Local Similarity 75.0%; Pred. No. 1e+03;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 QQMEQLRQ 12  
||:||||:  
Db 190 QQIEQLKQ 197

RESULT 40  
US-08-319-866-3  
; Sequence 3, Application US/08319866  
; Patent No. 5929223  
; GENERAL INFORMATION:  
; APPLICANT: Tully, Timothy P.  
; APPLICANT: Yin, Jerry C.  
; APPLICANT: Regulski, Michael  
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF GENES  
; TITLE OF INVENTION: ASSOCIATED WITH LONG-TERM MEMORY  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/319,866  
; FILING DATE: 7-OCT-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: CSHL94-03  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 861-6240  
; TELEFAX: (617) 861-9540  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 54 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-319-866-3

Query Match 55.4%; Score 31; DB 2; Length 54;  
Best Local Similarity 33.3%; Pred. No. 80;  
Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 KALVQQMEQLRQ 12  
||:||||:  
Db 40 KALIEELKSLKE 51

Search completed: September 8, 2003, 16:07:19  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2003, 16:06:18 ; Search time 25 Seconds  
(without alignments)  
65.906 Million cell updates/sec

Title: US-09-991-809-1  
Perfect score: 56  
Sequence: 1 KALVQOMEQLRQ 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513375 seqs, 137303645 residues

Total number of hits satisfying chosen parameters: 513375

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubppaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubppaa/US09C\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubppaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubppaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubppaa/US10C\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	12	11	US-09-991-809-1
2	56	100.0	396	9	US-09-800-729-207
3	56	100.0	396	10	US-09-987-107-33
4	56	100.0	401	10	US-09-987-107-36
5	56	100.0	429	10	US-09-987-107-34
6	51	91.1	396	11	US-09-802-640-16
7	49	87.5	391	9	US-09-800-729-208
8	49	87.5	395	10	US-09-987-107-35
9	47	83.9	382	9	US-09-800-729-206
10	47	83.9	382	10	US-09-987-107-37
11	46	82.1	10	10	US-09-791-378-561
12	46	82.1	10	10	US-09-826-290-317
13	46	82.1	10	10	US-09-826-290-333
14	46	82.1	10	10	US-09-826-290-424
15	46	82.1	10	11	US-09-791-393-226

16	46	82.1	10	11	US-09-791-389-226	Sequence 226, App
17	43	76.8	391	10	US-09-987-107-38	Sequence 38, Appl
18	43	76.8	391	12	US-10-316-253-238	Sequence 28, App
19	43	76.8	391	12	US-10-316-253-240	Sequence 240, App
20	43	76.8	391	12	US-10-316-253-242	Sequence 242, App
21	37	66.1	79	9	US-09-864-761-33972	Sequence 33972, A
22	36	64.3	380	11	US-09-782-816A-56	Sequence 56, Appl
23	36	64.3	616	15	US-10-103-313-460	Sequence 460, App
24	36	64.3	778	15	US-10-156-761-10767	Sequence 10767, A
25	36	64.3	1005	9	US-09-802-127-5	Sequence 5, Appl
26	35	62.5	234	9	US-09-867-550-890	Sequence 890, App
27	35	62.5	557	9	US-09-815-242-12165	Sequence 12165, A
28	35	62.5	557	9	US-09-815-242-12804	Sequence 12804, A
29	35	62.5	563	9	US-09-815-242-5464	Sequence 5464, Ap
30	35	62.5	566	16	US-10-282-287-4	Sequence 4, Appl
31	35	62.5	669	15	US-10-286-264-40	Sequence 40, Appl
32	35	62.5	669	16	US-10-278-536-12	Sequence 12, Appl
33	35	62.5	732	12	US-09-015-078-7	Sequence 7, Appl
34	35	62.5	732	12	US-09-015-078-8	Sequence 8, Appl
35	35	62.5	858	12	US-10-238-075-788	Sequence 788, App
36	35	62.5	2568	9	US-09-866-108-3	Sequence 3, Appl
37	34	60.7	284	9	US-09-745-763-207	Sequence 207, App
38	34	60.7	306	10	US-09-925-300-1668	Sequence 1668, Ap
39	34	60.7	368	11	US-09-533-029-8	Sequence 8, Appl
40	34	60.7	368	15	US-10-286-264-48	Sequence 48, Appl
41	34	60.7	368	15	US-10-295-403-152	Sequence 152, App
42	33	58.9	25	9	US-09-864-761-45458	Sequence 45458, A
43	33	58.9	284	12	US-10-137-870-118	Sequence 118, App
44	33	58.9	284	12	US-10-140-018-118	Sequence 118, App
45	33	58.9	284	12	US-10-140-021-118	Sequence 118, App

## ALIGNMENTS

RESULT 1  
US-09-991-809-1  
; Sequence 1, Application US/09991809  
; Publication No. US20030100014A1  
GENERAL INFORMATION:  
; APPLICANT: Jackowski, George  
; TITLE OF INVENTION: Apolipoprotein Biopolymer Markers Predictive of Type II Diabetes  
; FILE REFERENCE: 2132.111  
; CURRENT APPLICATION NUMBER: US/09/991,809  
; CURRENT FILING DATE: 2001-11-23  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 1  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-991-809-1

Query Match 100.0%; Score 56; DB 11; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0021;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVQOMEQLRQ 12  
DB 1 KALVQOMEQLRQ 12

RESULT 2  
US-09-800-729-207  
; Sequence 207, Application US/09800729  
; Patent No. US20020068319A1  
GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: F204491  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013

;; PRIOR FILING DATE: 2000-09-22  
;; PRIOR APPLICATION NUMBER: 60/155,709  
;; PRIOR FILING DATE: 1999-09-24  
;; NUMBER OF SEQ ID NOS: 217  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 207  
;; LENGTH: 396  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-800-729-207

Query Match 100.0%; Score 56; DB 9; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.076;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVQOMEQLRQ 12  
Db 316 KALVQOMEQLRQ 327

RESULT 3  
US-09-987-107-33  
; Sequence 33, Application US/09987107  
; Patent No. US20020156007A1  
; GENERAL INFORMATION:  
; APPLICANT: GRAVERSEN, Jonas  
; APPLICANT: MOESTRUP, Soren  
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES  
; FILE REFERENCE: GRAVERSENIA  
; CURRENT APPLICATION NUMBER: US/09/987,107  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/264,022  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: DK PA2001 00057  
; PRIOR FILING DATE: 2001-01-15  
; PRIOR APPLICATION NUMBER: DK PA2000 01682  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 33  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-987-107-33

Query Match 100.0%; Score 56; DB 10; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.076;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVQOMEQLRQ 12  
Db 316 KALVQOMEQLRQ 327

RESULT 4  
US-09-987-107-36  
; Sequence 36, Application US/09987107  
; Patent No. US20020156007A1  
; GENERAL INFORMATION:  
; APPLICANT: GRAVERSEN, Jonas  
; APPLICANT: MOESTRUP, Soren  
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES  
; FILE REFERENCE: GRAVERSENIA  
; CURRENT APPLICATION NUMBER: US/09/987,107  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/264,022  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: DK PA2001 00057  
; PRIOR FILING DATE: 2001-01-15  
; PRIOR APPLICATION NUMBER: DK PA2000 01682  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: PatentIn version 3.1

;; SEQ ID NO 36  
;; LENGTH: 401  
;; TYPE: PRT  
;; ORGANISM: Papio anubis  
US-09-987-107-36

Query Match 100.0%; Score 56; DB 10; Length 401;  
Best Local Similarity 100.0%; Pred. No. 0.077;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVQOMEQLRQ 12  
Db 300 KALVQOMEQLRQ 311

RESULT 5  
US-09-987-107-34  
; Sequence 34, Application US/09987107  
; Patent No. US20020156007A1  
; GENERAL INFORMATION:  
; APPLICANT: GRAVERSEN, Jonas  
; APPLICANT: MOESTRUP, Soren  
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES  
; FILE REFERENCE: GRAVERSENIA  
; CURRENT APPLICATION NUMBER: US/09/987,107  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/264,022  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: DK PA2001 00057  
; PRIOR FILING DATE: 2001-01-15  
; PRIOR APPLICATION NUMBER: DK PA2000 01682  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 34  
; LENGTH: 429  
; TYPE: PRT  
; ORGANISM: Macaca fascicularis  
US-09-987-107-34

Query Match 100.0%; Score 56; DB 10; Length 429;  
Best Local Similarity 100.0%; Pred. No. 0.082;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVQOMEQLRQ 12  
Db 316 KALVQOMEQLRQ 327

RESULT 6  
US-09-802-640-16  
; Sequence 16, Application US/09802640  
; Publication No. US20030036057A1  
; GENERAL INFORMATION:  
; APPLICANT: Braun, Andreas  
; APPLICANT: Bomsal Aruna  
; APPLICANT: Klevn Patrick  
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: 24736-2048  
; CURRENT APPLICATION NUMBER: US/09/802,640  
; CURRENT FILING DATE: 2001-03-09  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-802-640-16

Query Match 91.1%; Score 51; DB 11; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.53;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLR 11  
| | | | |  
Db 316 KALVOQMEQLR 326

RESULT 7  
US-09-800-729-208  
; Sequence 208, Application US/09800729  
; Patent No. US20020068319A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 208  
; LENGTH: 391  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-800-729-208

Query Match 87.5%; Score 49; DB 9; Length 391;  
Best Local Similarity 83.3%; Pred. No. 1.1;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLR 12  
| | | | |  
Db 316 KALVOQLEQFRQ 327

RESULT 8  
US-09-987-107-35  
; Sequence 35, Application US/09987107  
; Patent No. US20020156007A1  
; GENERAL INFORMATION:  
; APPLICANT: GRAVERSEN, Jonas  
; APPLICANT: MOESTRUP, Soren  
; TITLE OF INVENTION: APOLIPROTEINS ANALOGUES  
; FILE REFERENCE: GRAVERSENIA  
; CURRENT APPLICATION NUMBER: US/09/987,107  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/264,022  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: DK PA2001 00057  
; PRIOR FILING DATE: 2001-01-15  
; PRIOR APPLICATION NUMBER: DK PA2000 01682  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 35  
; LENGTH: 395  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-987-107-35

Query Match 87.5%; Score 49; DB 10; Length 395;  
Best Local Similarity 83.3%; Pred. No. 1.1;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLR 12  
| | | | |  
Db 316 KALVOQLEQFRQ 327

RESULT 9  
US-09-800-729-206

; Sequence 206, Application US/09800729  
; Patent No. US20020068319A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 206  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-800-729-206

Query Match 83.9%; Score 47; DB 9; Length 382;  
Best Local Similarity 83.3%; Pred. No. 2.4;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLR 12  
| | | | |  
Db 316 KALVOQVEDLRQ 327

RESULT 10  
US-09-987-107-37  
; Sequence 37, Application US/09987107  
; Patent No. US20020156007A1  
; GENERAL INFORMATION:  
; APPLICANT: GRAVERSEN, Jonas  
; APPLICANT: MOESTRUP, Soren  
; TITLE OF INVENTION: APOLIPROTEINS ANALOGUES  
; FILE REFERENCE: GRAVERSENIA  
; CURRENT APPLICATION NUMBER: US/09/987,107  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/264,022  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: DK PA2001 00057  
; PRIOR FILING DATE: 2001-01-15  
; PRIOR APPLICATION NUMBER: DK PA2000 01682  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 37  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Sus scrofa  
US-09-987-107-37

Query Match 83.9%; Score 47; DB 10; Length 382;  
Best Local Similarity 83.3%; Pred. No. 2.4;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLR 12  
| | | | |  
Db 316 KALVOQVEDLRQ 327

RESULT 11  
US-09-791-378-561  
; Sequence 561, Application US/09791378  
; Patent No. US20020142303A1  
; GENERAL INFORMATION:  
; APPLICANT: Parekh, Rajesh  
; TITLE OF INVENTION: PROTEINS, GENES AND THEIR USE FOR DIAGNOSIS AND TREATMENT OF  
; FILE REFERENCE: 9195-061-999  
; CURRENT APPLICATION NUMBER: US/09/791,378

```
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/750,395
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 677
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 561
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-378-561
```

```
Query Match      82.1%; Score 46; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2 ALVQOMEQLR 11
Db      1 ALVQOMEQLR 10
```

## RESULT 12

```
US-09-826-290-317
; Sequence 317 Application US/09826290
; Patent No. US20020164668A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Durham, L.Kathryn
; APPLICANT: Friedman, David L.
; APPLICANT: Herath, Herath Mudiyanselage Achula Chandrasiri
; APPLICANT: Kimmel, Lida H.
; APPLICANT: Patekh, Rajesh Bhikhu
; APPLICANT: Potter, David M.
; APPLICANT: Rohlf, Christian
; APPLICANT: Silber, B. Michael
; APPLICANT: Stiger, Thomas R.
; APPLICANT: Sunderland, P. Trey
; APPLICANT: Townsend, Robert Reid
; APPLICANT: White, Frost
```

```
; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
; TITLE OF INVENTION: Uses Therefor, Including Diagnosis and Treatment of
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 2572-1-001 N2
; CURRENT APPLICATION NUMBER: US/09/826,290
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 60/194,504
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/253,647
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 492
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 317
; LENGTH: 10
; TYPE: PRT
; ORGANISM: homo sapien
```

```
US-09-826-290-317
```

```
Query Match      82.1%; Score 46; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2 ALVQOMEQLR 11
Db      1 ALVQOMEQLR 10
```

## RESULT 13

```
US-09-826-290-333
; Sequence 333 Application US/09826290
; Patent No. US20020164668A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Durham, L.Kathryn
; APPLICANT: Friedman, David L.
; APPLICANT: Herath, Herath Mudiyanselage Achula Chandrasiri
```

```
; APPLICANT: Kimmel, Lida H.
; APPLICANT: Patekh, Rajesh Bhikhu
; APPLICANT: Potter, David M.
; APPLICANT: Rohlf, Christian
; APPLICANT: Silber, B. Michael
; APPLICANT: Stiger, Thomas R.
; APPLICANT: Sunderland, P. Trey
; APPLICANT: Townsend, Robert Reid
; APPLICANT: White, Frost
; APPLICANT: Williams, Stephen A.
; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
; TITLE OF INVENTION: Uses Therefor, Including Diagnosis and Treatment of
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 2572-1-001 N2
; CURRENT APPLICATION NUMBER: US/09/826,290
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 60/194,504
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/253,647
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 492
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 333
; LENGTH: 10
; TYPE: PRT
; ORGANISM: homo sapien
US-09-826-290-333
```

```
Query Match      82.1%; Score 46; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2 ALVQOMEQLR 11
Db      1 ALVQOMEQLR 10
```

## RESULT 14

```
US-09-826-290-424
; Sequence 424 Application US/09826290
; Patent No. US20020164668A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Durham, L.Kathryn
; APPLICANT: Friedman, David L.
; APPLICANT: Herath, Herath Mudiyanselage Achula Chandrasiri
; APPLICANT: Kimmel, Lida H.
; APPLICANT: Patekh, Rajesh Bhikhu
; APPLICANT: Potter, David M.
; APPLICANT: Rohlf, Christian
; APPLICANT: Silber, B. Michael
; APPLICANT: Stiger, Thomas R.
; APPLICANT: Sunderland, P. Trey
; APPLICANT: Townsend, Robert Reid
; APPLICANT: White, Frost
```

```
; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
; TITLE OF INVENTION: Uses Therefor, Including Diagnosis and Treatment of
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 2572-1-001 N2
; CURRENT APPLICATION NUMBER: US/09/826,290
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 60/194,504
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/253,647
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 492
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 424
; LENGTH: 10
; TYPE: PRT
; ORGANISM: homo sapien
```

```
US-09-826-290-424
```



Query Match 82.1%; Score 46; DB 10; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.086;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 ALVOQMEQLR 11  
|||  
Db 1 ALVOQMEQLR 10

## RESULT 15

US-09-791-393-226

; Sequence 226, Application US/09791393

; Publication No. US20030032200A1

; GENERAL INFORMATION:

; APPLICANT: Herath, Herath Mudiyanseelage Achula Chandrasiri

; APPLICANT: Parekh, Rajesh Bhikhu

; APPLICANT: Rohlf, Christian

; TITLE OF INVENTION: Proteins, Genes and Their Use for

; TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)

; FILE REFERENCE: 2543-1-001 N1

; CURRENT APPLICATION NUMBER: US/09/791,393

; CURRENT FILING DATE: 2002-01-02

; EARLIER APPLICATION NUMBER: GB 0004412.3

; EARLIER FILING DATE: 2000-02-24

; EARLIER APPLICATION NUMBER: GB 0030050.9

; EARLIER FILING DATE: 2000-12-08

; EARLIER APPLICATION NUMBER: US 60/254,830

; EARLIER FILING DATE: 2000-12-12

; NUMBER OF SEQ ID NOS: 308

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 226

; LENGTH: 10

; TYPE: PRT

; ORGANISM: homo sapien

US-09-791-393-226

Query Match 82.1%; Score 46; DB 11; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.086;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 ALVOQMEQLR 11  
|||  
Db 1 ALVOQMEQLR 10

## RESULT 16

US-09-791-389-226

; Sequence 226, Application US/09791389

; Publication No. US20030032773A1

; GENERAL INFORMATION:

; APPLICANT: Herath, Herath Mudiyanseelage Achula Chandrasiri

; APPLICANT: Parekh, Rajesh Bhikhu

; APPLICANT: Rohlf, Christian

; APPLICANT: Terrett, Jonathan Alexander

; APPLICANT: Tyson, Kerry Louise

; TITLE OF INVENTION: Proteins, Genes and Their Use for

; TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)

; FILE REFERENCE: 2543-1-001 N2

; CURRENT APPLICATION NUMBER: US/09/791,389

; CURRENT FILING DATE: 2001-02-23

; PRIOR APPLICATION NUMBER: GB 0004412.3

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: GB 0030050.9

; PRIOR FILING DATE: 2000-12-08

; PRIOR APPLICATION NUMBER: US 60/254,830

; PRIOR FILING DATE: 2000-12-12

; NUMBER OF SEQ ID NOS: 308

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 226

; LENGTH: 10

; TYPE: PRT

; ORGANISM: homo sapien  
US-09-791-389-226

Query Match 82.1%; Score 46; DB 11; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.086;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 ALVOQMEQLR 11  
|||  
Db 1 ALVOQMEQLR 10

## RESULT 17

US-09-987-107-38

; Sequence 38, Application US/09987107

; Patent No. US20020156007A1

; GENERAL INFORMATION:

; APPLICANT: GRAVERSEN, Jonas

; APPLICANT: MOESTRUP, Soren

; TITLE OF INVENTION: APOLIPROTEINS ANALOGUES

; FILE REFERENCE: GRAVERSENIA

; CURRENT APPLICATION NUMBER: US/09/987,107

; CURRENT FILING DATE: 2001-11-13

; PRIOR APPLICATION NUMBER: US 60/264,022

; PRIOR FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: DK PA2001 00057

; PRIOR FILING DATE: 2001-01-15

; PRIOR APPLICATION NUMBER: DK PA2000 01682

; PRIOR FILING DATE: 2000-11-10

; NUMBER OF SEQ ID NOS: 91

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 38

; LENGTH: 391

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-987-107-38

Query Match 76.8%; Score 43; DB 10; Length 391;  
Best Local Similarity 81.8%; Pred. No. 12;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 ALVOQMEQLRQ 12  
|||  
Db 317 ALVOQMEQLRQ 327

## RESULT 18

US-10-316-253-238

; Sequence 238, Application US/10316253

; Publication No. US20030162706A1

; GENERAL INFORMATION:

; APPLICANT: The Procter &amp; Gamble Company

; APPLICANT: Peters, Kevin

; APPLICANT: Thompson, Larry

; APPLICANT: Wang, Feng

; APPLICANT: Greis, Kenneth

; TITLE OF INVENTION: Angiogenesis Modulating Proteins

; FILE REFERENCE: 8865W

; CURRENT APPLICATION NUMBER: US/10/316,253

; CURRENT FILING DATE: 2002-12-10

; PRIOR APPLICATION NUMBER: US 60/355,295

; PRIOR FILING DATE: 2002-02-08

; NUMBER OF SEQ ID NOS: 308

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 238

; LENGTH: 391

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-10-316-253-238

Query Match 76.8%; Score 43; DB 12; Length 391;  
Best Local Similarity 81.8%; Pred. No. 12;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 ALVQOMEQRLQ 12  
|||:|  
Db 317 ALVQOMEQFRQ 327

RESULT 19  
US-10-316-253-240  
; Sequence 240, Application US/10316253  
; Publication No. US20030162706A1  
; GENERAL INFORMATION:  
; APPLICANT: The Procter & Gamble Company  
; APPLICANT: Peters, Kevin  
; APPLICANT: Thompson, Larry  
; APPLICANT: Wang, Feng  
; APPLICANT: Greis, Kenneth  
; TITLE OF INVENTION: Angiogenesis Modulating Proteins  
; FILE REFERENCE: 8865M  
; CURRENT APPLICATION NUMBER: US/10/316,253  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: US 60/355,295  
; PRIOR FILING DATE: 2002-02-08  
; NUMBER OF SEQ ID NOS: 308  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 240  
; LENGTH: 391  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-316-253-240

Query Match 76.8%; Score 43; DB 12; Length 391;  
Best Local Similarity 81.8%; Pred. No. 12;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 ALVQOMEQRLQ 12  
|||:|  
Db 317 ALVQOMEQFRQ 327

RESULT 20  
US-10-316-253-242  
; Sequence 242, Application US/10316253  
; Publication No. US20030162706A1  
; GENERAL INFORMATION:  
; APPLICANT: The Procter & Gamble Company  
; APPLICANT: Peters, Kevin  
; APPLICANT: Thompson, Larry  
; APPLICANT: Wang, Feng  
; APPLICANT: Greis, Kenneth  
; TITLE OF INVENTION: Angiogenesis Modulating Proteins  
; FILE REFERENCE: 8865M  
; CURRENT APPLICATION NUMBER: US/10/316,253  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: US 60/355,295  
; PRIOR FILING DATE: 2002-02-08  
; NUMBER OF SEQ ID NOS: 308  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 242  
; LENGTH: 391  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-316-253-242

Query Match 76.8%; Score 43; DB 12; Length 391;  
Best Local Similarity 81.8%; Pred. No. 12;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 ALVQOMEQRLQ 12  
|||:|  
Db 317 ALVQOMEQFRQ 327

RESULT 21

US-09-864-761-33972  
; Sequence 33972, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wenheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecm1ca-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26;  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 33972  
; LENGTH: 79  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC007050.23  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 25  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 18  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.4  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 33  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 39  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 12  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 21  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 34  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 24  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 24  
US-09-864-761-33972

Query Match 66.1%; Score 37; DB 9; Length 79;  
Best Local Similarity 58.3%; Pred. No. 23;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 KALVOQMEQLRQ 12  
Db 50 QAVVOQOQLOO 61

## RESULT 22

US-09-782-816A-56  
; Sequence 56, Application US/09782816A  
; Publication No. US20030032711A1  
; GENERAL INFORMATION:  
; APPLICANT: Sharp, David J.  
; APPLICANT: Rogers, Gregory C.  
; APPLICANT: Scholey, Jonathan M.  
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR  
; FILE REFERENCE: UC069, 001A  
; CURRENT APPLICATION NUMBER: US/09/782, 816A  
; CURRENT FILING DATE: 2001-02-14  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 56  
; LENGTH: 380  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-782-816A-56

Query Match 64.3%; Score 36; DB 11; Length 380;  
Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 KALVOQMEQLRQ 12  
Db 169 KALISQVEEPKQ 180

## RESULT 23

US-10-103-313-460  
; Sequence 460, Application US/10103313  
; Publication No. US20030082758A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P1207C1  
; CURRENT APPLICATION NUMBER: US/10/103, 313  
; CURRENT FILING DATE: 2002-03-12  
; NUMBER OF SEQ ID NOS: 653  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 460  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-103-313-460

Query Match 64.3%; Score 36; DB 15; Length 616;  
Best Local Similarity 63.6%; Pred. No. 2.8e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KALVOQMEQLR 11  
Db 479 RALTEQOEQLR 489

## RESULT 24

US-10-156-761-10767  
; Sequence 10767, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156, 761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 10767  
; LENGTH: 778  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-10767

Query Match 64.3%; Score 36; DB 15; Length 778;  
Best Local Similarity 70.0%; Pred. No. 3.6e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 LVQOMEQLRQ 12  
Db 672 LVQOEQLR 681

## RESULT 25

US-09-802-127-5  
; Sequence 5, Application US/09802127  
; Patent No. US20020045212A1  
; GENERAL INFORMATION:  
; APPLICANT: Glucksmann, Maria Alexandra  
; APPLICANT: Meyers, Rachel  
; TITLE OF INVENTION: NO. US20020045212A1 Human GTPase Activator Proteins  
; FILE REFERENCE: 035800/158994  
; CURRENT APPLICATION NUMBER: US/09/802, 127  
; CURRENT FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: 60/185, 611  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 1005  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-802-127-5

Query Match 64.3%; Score 36; DB 9; Length 1005;  
Best Local Similarity 63.6%; Pred. No. 4.7e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KALVOQMEQLR 11  
Db 914 RALTEQOEQLR 924

## RESULT 26

US-09-867-550-890  
; Sequence 890, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Foad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: NO. US20020082206A1 Polynucleotides from Atherogenic Cells and  
; FILE REFERENCE: 21402-013 (Cite-313)  
; CURRENT APPLICATION NUMBER: US/09/867, 550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208, 427  
; PRIOR FILING DATE: 2000-05-30

```
/ NUMBER OF SEQ ID NOS: 2125
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 890
/ LENGTH: 234
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: VARIANT
/ LOCATION: (1)
/ OTHER INFORMATION: wherein Xaa may be any one of Phe or Ser or Tyr or Cys or Leu or
/ OTHER INFORMATION: His or Arg or Ile or Thr or Asn or Val or Ala or Asp or Gly
US-09-867-550-890
```

```
Query Match          62.5%; Score 35; DB 9; Length 234;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 KALVQMEQ 9
        |||:||||:
Db      199 KALVQMEK 207
```

```
RESULT 27
US-09-815-242-12165
/ Sequence 12165, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyckind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR FILING DATE: 2001-02-16
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 12165
/ LENGTH: 557
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
US-09-815-242-12165
```

```
Query Match          62.5%; Score 35; DB 9; Length 557;
Best Local Similarity 70.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      3 LVQQMEQLRQ 12
        |||:||||:
Db      367 LLQRMQLRQ 376
```

```
RESULT 28
US-09-815-242-12804
```

```
/ Sequence 12804, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyckind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR FILING DATE: 2001-02-16
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 12804
/ LENGTH: 557
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
US-09-815-242-12804
```

```
Query Match          62.5%; Score 35; DB 9; Length 557;
Best Local Similarity 70.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      3 LVQQMEQLRQ 12
        |||:||||:
Db      367 LLQRMQLRQ 376
```

```
RESULT 29
US-09-815-242-5464
/ Sequence 5464, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyckind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR FILING DATE: 2000-10-23
```

```

; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 5464
; LENGTH: 563
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-5464

```

```

Query Match      62.5%; Score 35; DB 9; Length 563;
Best Local Similarity 70.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      3 LVQOMEQLRQ 12
      ||:|||||
Db      375 LLQRMQLRQ 384

```

```

RESULT 30
; US-10-282-287-4
; Sequence 4, Application US/10282287
; Publication No. US20030129633A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Zhang, Dan
; APPLICANT: Whipple, Richard
; TITLE OF INVENTION: DNA REPLICATION PROTEINS OF GRAM POSITIVE BACTERIA AND
; FILE REFERENCE: 22221/1002
; CURRENT APPLICATION NUMBER: US/10/282,287
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US/09/235,245
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/074,522
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 60/093,727
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 4
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-10-282-287-4

```

```

Query Match      62.5%; Score 35; DB 16; Length 566;
Best Local Similarity 70.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      3 LVQOMEQLRQ 12
      ||:|||||
Db      375 LLQRMQLRQ 384

```

```

RESULT 31
; US-10-286-264-40
; Sequence 40, Application US/10286264
; Publication No. US20030093837A1
; GENERAL INFORMATION:
; APPLICANT: Keddie, James
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Zhang, James
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Pineda, Omaira
; APPLICANT: Heard, Jacqueline
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Adam, Luc

```

```

; APPLICANT: Broun, Pierre
; APPLICANT: Reuber, Lynne
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Samaha, Raymond
; TITLE OF INVENTION: POLYNUCLEOTIDES FOR SEED TRAIT ALTERATION
; FILE REFERENCE: MBI-008
; CURRENT APPLICATION NUMBER: US/10/286,264
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 40
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G634
; US-10-286-264-40

```

```

Query Match      62.5%; Score 35; DB 15; Length 669;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 KALVQOMEQLRQ 12
      :||:||||:
Db      591 RELIQQQQLRQ 602

```

```

RESULT 32
; US-10-278-536-12
; Sequence 12, Application US/10278536
; Publication No. US2003013186A1
; GENERAL INFORMATION:
; APPLICANT: Samaha, Raymond
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Keddie, James
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Adam, Luc
; APPLICANT: Broun, Pierre
; TITLE OF INVENTION: STRESS-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-011
; CURRENT APPLICATION NUMBER: US/10/278,536
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 12
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G634
; US-10-278-536-12

```

```

Query Match      62.5%; Score 35; DB 16; Length 669;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 KALVQOMEQLRQ 12
      :||:||||:
Db      591 RELIQQQQLRQ 602

```

```

RESULT 33
; US-09-015-078-7

```

; Sequence 7, Application US/09015078  
; Publication No. US20030152579A1  
; GENERAL INFORMATION:  
; APPLICANT: Suebaum, Sebastian  
; Labigne, Agnes  
; TITLE OF INVENTION: Cloning and Characterization of the flba  
; Gene of H. Pylori, Production of Aflagellate Strains  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fimegan, Henderson, Farabow, Garrett &  
; Dunner  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/015,078  
; FILING DATE: 29-Jan-1998  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 02356.0073-01000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 408-4000  
; TELEFAX: (202) 408-4400  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 732 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; \* MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-015-078-7  
Query Match 62.5%; Score 35; DB 12; Length 732;  
Best Local Similarity 72.7%; Pred. No. 5e+02;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 KALVQMEQLR 11  
DB 695 KALSNMEQAR 705  
RESULT 34  
US-09-015-078-8  
; Sequence 8, Application US/09015078  
; Publication No. US20030152579A1  
; GENERAL INFORMATION:  
; APPLICANT: Suebaum, Sebastian  
; Labigne, Agnes  
; TITLE OF INVENTION: Cloning and Characterization of the flba  
; Gene of H. Pylori, Production of Aflagellate Strains  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fimegan, Henderson, Farabow, Garrett &  
; Dunner  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/015,078  
; FILING DATE: 29-Jan-1998  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 02356.0073-01000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 408-4000  
; TELEFAX: (202) 408-4400  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 732 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-015-078-8  
Query Match 62.5%; Score 35; DB 12; Length 732;  
Best Local Similarity 72.7%; Pred. No. 5e+02;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 KALVQMEQLR 11  
DB 695 KALSNMEQAR 705  
RESULT 35  
US-10-238-075-788  
; Sequence 788, Application US/10238075  
; Publication No. US20030148324A1  
; GENERAL INFORMATION:  
; APPLICANT: I.N.S.E.R.M.  
; TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolat  
; FILE REFERENCE: BLANDINE  
; CURRENT APPLICATION NUMBER: US/10/238.075  
; PRIOR FILING DATE: 2002-09-10  
; PRIOR APPLICATION NUMBER: 0003145  
; PRIOR FILING DATE: 2000-03-10  
; NUMBER OF SEQ ID NOS: 1576  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 788  
; LENGTH: 858  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-238-075-788  
Query Match 62.5%; Score 35; DB 12; Length 858;  
Best Local Similarity 66.7%; Pred. No. 5.8e+02;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 KALVQMEQLRQ 12  
DB 486 KALAQOIMTLRQ 497  
RESULT 36  
US-09-866-108-3  
; Sequence 3, Application US/09866108  
; Patent No. US20020048800A1  
; GENERAL INFORMATION:  
; APPLICANT: GU, Yizhong  
; APPLICANT: JI, Yonggang  
; APPLICANT: PENN, Sharon G.  
; APPLICANT: HANZEL, David K.  
; APPLICANT: RANK, David R.  
; APPLICANT: CHEN, Wensheng  
; APPLICANT: SHANNON, Mark

TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE  
FILE REFERENCE: AECOMICA-7  
CURRENT APPLICATION NUMBER: US/09/866,108  
CURRENT FILING DATE: 2001-05-25  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 60/266,860  
PRIOR FILING DATE: 2001-02-05  
NUMBER OF SEQ ID NOS: 15752  
SOFTWARE: Aecomica Sequence Listing Engine  
SEQ ID NO 3  
LENGTH: 2568  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-866-108-3

Query Match 62.5%; Score 35; DB 9; Length 2568;  
Best Local Similarity 66.7%; Pred. No. 1.8e+03;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 VQOMEQ 12  
Db 1712 IKOLEQLRQ 1720

RESULT 37  
US-09-745-763-207  
Sequence 207, Application US/09745763  
Patent No. US20020065394A1  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
McCoy, John M.  
Lavallie, Edward R.  
Collins-Racie, Lisa A.  
Evans, Cheryl  
Merberg, David  
Treacy, Maurice  
Spaulding, Vikki  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM  
NUMBER OF SEQUENCES: 219  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.

ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/745,763  
FILING DATE: 18-Jun-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Sprunger, Suzanne A.  
REGISTRATION NUMBER: 41,323  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8284  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 207:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 284 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 207:  
US-09-745-763-207

Query Match 60.7%; Score 34; DB 9; Length 284;  
Best Local Similarity 77.8%; Pred. No. 2.8e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KALVQMEQ 9  
Db 121 KQIVQMLEQ 129

RESULT 38  
US-09-925-300-1668  
Sequence 1668, Application US/09925300  
Patent No. US20020151681A1  
GENERAL INFORMATION:  
APPLICANT: Craig Ruben,  
Steve Ruben  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA101  
CURRENT APPLICATION NUMBER: US/09/925,300  
PRIOR FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05988  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 1890  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 1668  
LENGTH: 306  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-925-300-1668

Query Match 60.7%; Score 34; DB 10; Length 306;  
Best Local Similarity 77.8%; Pred. No. 3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KALVQMEQ 9  
Db 143 KQIVQMLEQ 151

RESULT 39  
US-09-533-029-8  
Sequence 8, Application US/09533029  
Publication No. US20030046723A1  
GENERAL INFORMATION:  
APPLICANT: Heard, Jacqueline

```

; APPLICANT: Brown, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddle, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 8
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G629
US-09-533-029-8
```

```

Query Match      60.7%; Score 34; DB 11; Length 368;
Best Local Similarity 63.6%; Pred. No. 3.6e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 KALVQOMEQLR 11
        |||||:|:|
Db      113 KAVYQGLEESR 123
```

```

RESULT 40
US-10-286-264-48
; Sequence 48, Application US/10286264
; Publication No. US20030093837A1
; GENERAL INFORMATION:
; APPLICANT: Keddle, James
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Zhang, James
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Pineda, Omaira
; APPLICANT: Heard, Jacqueline
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Adam, Luc
; APPLICANT: Brown, Pierre
; APPLICANT: Reuber, Lynne
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Samaha, Raymond
; TITLE OF INVENTION: POLYNUCLEOTIDES FOR SEED TRAIT ALTERATION
; FILE REFERENCE: MBI-008
; CURRENT APPLICATION NUMBER: US/10/286,264
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 48
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G629
US-10-286-264-48
```

```

Query Match      60.7%; Score 34; DB 15; Length 368;
Best Local Similarity 63.6%; Pred. No. 3.6e+02;
```

```

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy      1 KALVQOMEQLR 11
        |||||:|:|
Db      113 KAVYQGLEESR 123
```

```

Search completed: September 8, 2003, 16:11:24
Job time : 26 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2003, 16:03:33 ; Search time 21 Seconds  
(without alignments)  
54.954 Million cell updates/sec

Title: US-09-991-809-1  
Perfect score: 56  
Sequence: 1 KALVQWMEQLRQ 12

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir16:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	100.0	396	1 LPHU4	apolipoprotein A-I
2	56	100.0	401	2 A47141	apolipoprotein A-I
3	56	100.0	429	2 S29565	apolipoprotein A-I
4	51	91.1	91	2 I54248	apolipoprotein A-I
5	49	87.5	391	2 B40892	apolipoprotein A-I
6	49	87.5	395	2 A40892	apolipoprotein A-I
7	49	87.5	399	2 A40892	apolipoprotein A-I
8	44	78.6	394	2 A25281	apolipoprotein A-I
9	43	76.8	391	1 LPHU4	apolipoprotein A-I
10	39	69.6	311	2 A56235	apolipoprotein A-I
11	39	69.6	323	2 I49529	apolipoprotein A-I
12	39	69.6	2712	2 T05113	transcription factor
13	38	67.9	219	2 B41886	hypothetical prote
14	38	67.9	597	2 S40998	flagellar membrane
15	38	67.9	893	2 G88551	hypothetical prote
16	37	66.1	177	2 B69520	protein M01A8.2 [i
17	37	66.1	359	2 A12368	hypothetical prote
18	37	66.1	893	2 A12368	hypothetical prote
19	36	64.3	208	2 E75520	toxin secretion AB
20	36	64.3	370	2 AG0761	hypothetical prote
21	36	64.3	498	2 H83880	probable propionol
22	36	64.3	1189	2 A54817	hypothetical prote
23	36	64.3	1279	2 T13613	ATPase ScII, chrom
24	36	64.3	1738	2 T14867	hypothetical prote
25	36	64.3	2024	2 A54103	interaplin - slime
26	35	62.5	79	2 AE0012	centrosome autocat
27	35	62.5	190	2 H71245	conserved hypotnet
28	35	62.5	338	2 S73441	hypothetical prote
29	35	62.5	394	2 E64028	MG032 homolog B01_
					hypothetical prote

30	35	62.5	461	2 S77437	argininosuccinate
31	35	62.5	531	2 A84444	hypothetical prote
32	35	62.5	565	2 G89813	DNA polymerase III
33	35	62.5	633	2 T40124	kinesin-like motor
34	35	62.5	733	2 E71937	flagellar biosynth
35	35	62.5	733	2 A64650	flagellar biosynth
36	35	62.5	918	2 C96829	unknown protein P1
37	35	62.5	1043	2 A11906	hypothetical prote
38	35	62.5	1413	2 G86301	hypothetical prote
39	35	62.5	4845	2 T31067	probable retroelem
40	34	60.7	124	2 AF0088	BIR repeat contain
41	34	60.7	213	2 C64041	hypothetical prote
42	34	60.7	225	2 AD2429	hypothetical prote
43	34	60.7	235	2 AF3610	transcription regu
44	34	60.7	262	2 AD0164	conserved hypotnet
45	34	60.7	368	2 AC2201	transcription regu

## ALIGNMENTS

RESULT 1  
LPHU4  
apolipoprotein A-IV precursor [validated] - human  
N:Alternate names: apoA-IV  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1987 #sequence revision 30-Sep-1987 #text change 08-Dec-2000  
C:Accession: A94137; A94059; A24449; A29330; A26280; I37177; C54223; A61203; A26481; S02  
R:Karathanasis, S.K.; Oeltgen, P.; Haddad, I.A.; Antonarakis, S.E.  
Proc. Natl. Acad. Sci. U.S.A. 83, 8457-8461, 1986  
A:Title: Structure, evolution, and polymorphisms of the human apolipoprotein A4 gene (AF  
A:Reference number: A94137; MUID:87041474; PMID:3095836  
A:Accession: A94137  
A:Molecule type: DNA  
A:Residues: 1-396 <KAR1>  
A:Cross-references: GB:M14642; NID:G178760; PIDN:AAA51745.1; PID:G178761  
R:Karathanasis, S.K.  
Proc. Natl. Acad. Sci. U.S.A. 82, 6374-6378, 1985  
A:Title: Apolipoprotein multigene family: tandem organization of human apolipoprotein AI  
A:Reference number: A94059; MUID:86016704; PMID:3931073  
A:Accession: A94059  
A:Molecule type: mRNA  
A:Residues: 135-378 <KAR2>  
A:Cross-references: GB:M10373; NID:G563319; PIDN:AA59516.1; PID:G563320  
R:Karathanasis, S.K.; Yunis, J.  
Biochemistry 25, 3962-3970, 1986  
A:Title: Structure, evolution, and tissue-specific synthesis of human apolipoprotein AIV  
A:Reference number: A24449; MUID:86296629; PMID:3755616  
A:Accession: A24449  
A:Molecule type: mRNA  
A:Residues: 1-396 <KAR3>  
A:Cross-references: GB:M13654; NID:G178758; PIDN:AAA51744.1; PID:G178759  
R:Gordon, J.I.; Bisgaler, C.L.; Sims, H.F.; Sachdev, O.P.; Glickman, R.M.; Strauss, A.W.  
J. Biol. Chem. 259, 4684-474, 1984  
A:Title: Biosynthesis of human preapolipoprotein A-IV.  
A:Reference number: A92475; MUID:84161950; PMID:6706947  
A:Contents: annotation; signal sequence cleavage site  
R:Elshourbagy, N.A.; Walker, D.W.; Boguski, M.S.; Freeman, M.; Gordon, J.I.;  
J. Biol. Chem. 262, 7973-7981, 1987  
A:Title: Structure and expression of the human apolipoprotein A-IV gene.  
A:Reference number: A29330; MUID:87250378; PMID:3036793  
A:Accession: A29330  
A:Molecule type: DNA  
A:Residues: 1-157, 'T', 159, 'Y', 161-378, 'R', 280-326, 'T', 328-379, 'H', 381-396 <ELS>  
A:Cross-references: GB:J02758; NID:G178756; PIDN:AAA96731.1; PID:G178757  
R:Elshourbagy, N.A.; Walker, D.W.; Boguski, M.S.; Gordon, J.I.; Taylor, J.M.  
J. Biol. Chem. 261, 1998-2002, 1986  
A:Title: The nucleotide and derived amino acid sequence of human apolipoprotein A-IV mRNA  
A:Reference number: A26280; MUID:86111885; PMID:3080432  
A:Accession: A26280  
A:Molecule type: mRNA  
A:Residues: 21-157, 'T', 159, 'Y', 161-278, 'R', 280-326, 'T', 328-379, 'H', 381-396 <EL2>  
A:Cross-references: GB:M14566; NID:G178778; PIDN:AAA51748.1; PID:G178779

R:Yang, C.Y.; Gu, Z.W.; Chong, I.S.; Xiong, W.J.; Rosseneu, M.; Yang, H.X.; Lee, B.R.; G  
Biochim. Biophys. Acta 1002, 231-237, 1989  
A:Title: The primary structure of human apolipoprotein A-IV.  
A:Reference number: 137177; MUID:89194198; PMID:2930771  
A:Accession: 137177  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-157, 'T', 159, 'Y', 161-278, 'R', 280-396 <YANI>  
A:Cross-references: EMBL:X13629; NID:G28761; PIDs:CAA31955.1; PID:g28762  
A:Note: submitted to the EMBL Data Library, January 1989  
R:Kunitake, S.T.; Carilli, C.T.; Lau, K.; Procter, A.A.; Naya-Vigne, J.; Kane, J.P.  
Biochemistry 33, 1988-1993, 1994  
A:Title: Identification of proteins associated with apolipoprotein A-I-containing lipop  
A:Reference number: A54223; MUID:94162201; PMID:8117655  
A:Accession: C54223  
A:Molecule type: protein  
A:Residues: 'X', 22, 'X', 24, 'X', 26-31, 'X', 33-34 <KUN>  
R:Tenkainen, H.; Lukka, M.; Jauhainen, M.; Metsa, J.; Baumann, M.; Peltonen, L.; Ehnholm  
Arterioscler. Thromb. 11, 851-856, 1991  
A:Title: The mutation causing the common apolipoprotein A-IV polymorphism is a glutamine  
A:Reference number: A61203; MUID:91291788; PMID:2065039  
A:Accession: A61203  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 378-379, 'H', 381-382 <TEN>  
C:Comment: ApoA-IV is a major lipoprotein of lymph chylomicrons. In human plasma, it is  
C:Genetics: ApoA-IV is synthesized primarily in the intestine.  
A:Gene: GDB:APOA4  
A:Cross-references: GDB:119000; OMIM:107690  
A:Map position: 11q23-11q23  
A:Introns: 17/1, 59/2  
C:Superfamily: apolipoprotein A-I  
C:Keywords: chylomicron; HDL; intestine; lipid binding; lipid transport; lipoprotein; pl  
F:1-20/Doman: signal sequence; status experimental <SID>  
F:21-396/Product: apolipoprotein A-IV; status experimental <MAT>  
Query Match 100.0%; Score 56; DB 1; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.037;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KALVOQMEQLRQ 12  
DB 316 KALVOQMEQLRQ 327  
RESULT 2  
A:7141  
apolipoprotein A-IV I isoform - baboon (fragment)  
C:Species: Papio sp. (baboon)  
C>Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 12-Apr-1995  
C:Accession: A47141  
R:Hixson, J.E.; Kammerer, C.M.; Mott, G.E.; Britten, M.L.; Birnbaum, S.; Powers, P.K.; V  
J. Biol. Chem. 268, 15667-15673, 1993  
A:Title: Baboon apolipoprotein A-IV. Identification of lys76--Glu that distinguishes w  
A:Reference number: A47141; MUID:93340170; PMID:8101842  
A:Accession: A47141  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-401 <HIX>  
A:Experimental source: intestine  
A:Note: sequence extracted from NCBI backbone (NCBIN:136009, NCBI:P:136010)  
C:Superfamily: apolipoprotein A-I  
Query Match 100.0%; Score 56; DB 2; Length 401;  
Best Local Similarity 100.0%; Pred. No. 0.037;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KALVOQMEQLRQ 12  
DB 300 KALVOQMEQLRQ 311

RESULT 3  
S29565  
apolipoprotein A-IV - crab-eating macaque  
C:Species: Macaca fascicularis (Crab-eating macaque)  
C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 13-Aug-1999  
C:Accession: S30195; S29565  
R:Osada, J.; Pocovi, M.; Nicolosi, R.J.; Schaefer, E.J.; Ordovas, J.M.  
Biochim. Biophys. Acta 1172, 335-339, 1993  
A:Title: Nucleotide sequences of the Macaca fascicularis apolipoprotein C-III and A-IV g  
A:Reference number: S30195; MUID:93192330; PMID:8448212  
A:Accession: S30195  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-429 <OSA>  
A:Cross-references: EMBL:X68361; NID:G38050; PIDs:CAA48421.1; PID:g38051  
C:Genetics: 17/1; 59/2  
C:Superfamily: apolipoprotein A-I  
Query Match 100.0%; Score 56; DB 2; Length 429;  
Best Local Similarity 100.0%; Pred. No. 0.04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KALVOQMEQLRQ 12  
DB 316 KALVOQMEQLRQ 327  
RESULT 4  
S4248  
apolipoprotein A-IV - chimpanzee (fragment)  
C:Species: Pan troglodytes (chimpanzee)  
C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 13-Aug-1999  
C:Accession: S4248  
R:Kamboh, M.I.; Kelly, L.J.; Ahn, Y.I.; Ferrell, R.E.  
Hum. Biol. 66, 625-638, 1994  
A:Title: Genetic polymorphism of apolipoprotein A-IV in the chimpanzee: common deletion  
A:Reference number: S4248; MUID:94374966; PMID:8088751  
A:Accession: S4248  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-91 <RES>  
A:Cross-references: GB:S72300; NID:G601800; PIDs:AAB31973.1; PID:g601801  
C:Genetics: APOA4  
A:Gene: APOA4  
C:Superfamily: apolipoprotein A-I  
Query Match 91.1%; Score 51; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 0.06;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KALVOQMEQLR 11  
DB 11 KALVOQMEQLR 21  
RESULT 5  
B40892  
apolipoprotein A-IV precursor - mouse (strain 129)  
C:Species: Mus musculus (house mouse)  
C>Date: 27-Mar-1992 #sequence\_revision 10-Apr-1992 #text\_change 13-Aug-1999  
C:Accession: B40892  
R:Reue, K.; Leete, T.H.  
J. Biol. Chem. 266, 12715-12721, 1991  
A:Title: Genetic variation in mouse apolipoprotein A-IV due to insertion and deletion in  
A:Reference number: A40892; MUID:91286309; PMID:1648102  
A:Accession: B40892  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-391 <REU>  
A:Cross-references: GB:M64248; NID:9191884; PIDs:AAA37214.1; PID:g191885  
A:Note: the authors translated the codon CTG for residue 87 as Glu, GAG for residue 207  
C:Superfamily: apolipoprotein A-I



```

C:Accession: A56235
R:Katsuma, K.; Fujiwara, K.T.; Noda, M.; Nishizawa, M.
Mol. Cell. Biol. 14, 7581-7591, 1994
A:Title: MafB, a new Maf family transcription activator that can associate with Maf and
A:Reference number: A56235; MUID:95021288; PMID:7935473
A:Accession: A56235
A:Status: preliminary
A:Molecule type: DNA
A:Cross-references: GB:D28600; NID:9516723; PIDN:BAA05938.1; PID:9516724
C:Genetics:
A:Introns: #status absent
C:Superfamily: maf transforming protein; maf homology
C:Keywords: DNA binding; homodimer; leucine zipper
F:200-289/Domain: maf homology <MAF>

Query Match          69.6%; Score 39; DB 2; Length 311;
Best Local Similarity 70.0%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LVQVQMEQLRQ 12
Db 261 LIQVVEQLKQ 270

RESULT 11
149529
transcription factor-kr - mouse
C:Species: Mus musculus (house mouse)
C:Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 16-Jul-1999
A:Accession: 149529
R:Cordeiro, S.P.; Barsh, G.S.
Cell 79, 1025-1034, 1994
A:Title: The mouse segmentation gene kr encodes a novel basic domain-leucine zipper tran
A:Reference number: A55200; MUID:95094266; PMID:8001130
A:Accession: 149529
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-323 <RBS>
A:Cross-references: GB:L36435; NID:9625043; PIDN:AAA65689.1; PID:9625044
C:Superfamily: maf transforming protein; maf homology
C:Keywords: leucine zipper; transcription factor
F:212-301/Domain: maf homology <MAF>

Query Match          69.6%; Score 39; DB 2; Length 323;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LVQVQMEQLRQ 12
Db 273 LIQVVEQLKQ 282

RESULT 12
T05113
hypothetical protein F28M20.240 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
A:Accession: T05113
R:Bevan, M.; Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Hohnes, J.; Mewes,
submitted to the Protein Sequence Database, November 1998
A:Reference number: Z15398
A:Accession: T05113
A:Molecule type: DNA
A:Residues: 1-2712 <BR>
A:Cross-references: EMBL:AL031004
A:Experimental source: cultivar Columbia; BAC clone F28M20
C:Genetics:
A:Map position: 4
A:Introns: 17/3; 240/1; 1950/1; 2118/3; 2381/1; 2599/3; 2645/1; 2679/2
A>Note: F28M20.240

Query Match          69.6%; Score 39; DB 2; Length 2712;

```

```

Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KALVQMEQLRQ 12
Db 1873 KALVQMRSLKQ 1884

RESULT 13
B1886
flagellar membrane protein fliz - Bacillus subtilis
N:Alternate names: flagellar formation protein fliz
C:Species: Bacillus subtilis
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999
A:Accession: B1886; H69625
R:Bischoff, D.S.; Weinreich, M.D.; Ordal, G.W.
J. Bacteriol. 174, 4017-4025, 1992
A:Title: Nucleotide sequences of Bacillus subtilis flagellar biosynthetic genes fliz and
A:Reference number: A1886; MUID:92283757; PMID:1597417
A:Accession: B1886
A:Molecule type: DNA
A:Residues: 1-219 <BIS>
A:Cross-references: GB:M87005; NID:9142928; PIDN:AAA2452.1; PID:9142930
A>Note: sequence extracted from NCBI database (NCBI:106465, NCBI:106467)
R:Kunst, F.; Ogawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bexter,
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funai, S.; Galizzi, A.; Galler,
ich, J.; Harwood, C.R.; Henauc, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lartinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.W.; Levine, A.; Liu, H.; Masuda, S.; Mausel,
Y. M.; Ogawa, K.; Ogilawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpe, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Dancin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A6580; MUID:98044033; PMID:9384377
A:Accession: H69625
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-219 <KUN>
A:Cross-references: GB:Z59112; GB:AL009126; NID:92633902; PIDN:CAB13507.1; PID:el185225;
A:Experimental source: strain 168
C:Genetics:
A:Gene: fliz
C:Keywords: transmembrane protein

Query Match          67.9%; Score 38; DB 2; Length 219;
Best Local Similarity 58.3%; Pred. No. 25;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KALVQMEQLRQ 12
Db 194 KALVQMEQLKQ 205

RESULT 14
S40998
hypothetical protein M01A8.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
A:Accession: S40998
R:Hawkins, T.; Thomas, K.
submitted to the EMBL Data Library, October 1993
A:Reference number: S40997
A:Accession: S40998
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-597 <HAM>
A:Cross-references: EMBL:Z27081

```

C:Genetics:  
A:introns: 40/3; 76/1; 125/1; 199/1; 240/1; 274/1; 292/3; 392/3; 429/2; 486/2; 521/1

Query Match  
Best Local Similarity 66.7%; Score 38; DB 2; Length 597;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KALVQOMEOLRQ 12  
Db 429 KILVQIJDRLR 440

RESULT 15  
G88551  
protein M01A8.2 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: G88551  
R:anonymous, The C. elegans Sequencing Consortium.  
S:Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.eleg  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: G88551  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-893 <STO>  
A:Cross-references: GB:chr\_III; PIDN:CAAB1607.1; PID:G3878558; GSPDB:GN00021; CESP:M01A8  
C:Genetics:  
A:Gene: M01A8.2  
A:Map position: 3

Query Match  
Best Local Similarity 67.9%; Score 38; DB 2; Length 893;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KALVQOMEOLRQ 12  
Db 725 KILVQIJDRLR 736

RESULT 16  
B69520  
hypoetical protein AF2162 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C:Accession: B69520  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirschner, E.F.;  
Glodek, A.; Zhou, L.; Overbeek, R.; Goeyne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: B69520  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-177 <KLE>  
A:Cross-references: GB:AE000955; GB:AE000782; NID:G2689278; PIDN:AA89098.1; PID:G264837  
C:Superfamily: Archaeoglobus fulgidus hypothetical protein AF2162

Query Match  
Best Local Similarity 66.1%; Score 37; DB 2; Length 177;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KALVQOMEOLRQ 12  
Db 153 KALVQIJDRLR 164

RESULT 17

A12368  
hypoetical protein alr4505 [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: A12368  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriugu  
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: A12368  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-359 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA876204.1; PID:G17133641; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: alr4505

Query Match  
Best Local Similarity 66.1%; Score 37; DB 2; Length 359;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KALVQOMEOLRQ 12  
Db 210 KALVQIJDRLR 221

RESULT 18  
AH2007  
toxin secretion ABC transporter ATP-binding protein alr1614 [imported] - Nostoc sp. (str  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AH2007  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriugu  
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AH2007  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-893 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA877980.1; PID:G17133643; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: alr1614

Query Match  
Best Local Similarity 66.1%; Score 37; DB 2; Length 893;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALVQOMEOLRQ 12  
Db 437 ARVQELQIRQ 447

RESULT 19  
E75520  
hypoetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: E75520  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: E75520  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-208 <WHI>  
A:Cross-references: GB:AE001902; GB:AE000513; NID:g6458103; PIDN:AAF10011.1; PID:g645811  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0423  
A:Map position: 1

Query Match 64.3%; Score 36; DB 2; Length 208;  
Best Local Similarity 88.9%; Pred. No. 53;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVQMEQLR 11  
Db 157 LVQMEQLR 165

## RESULT 20

AG0761  
probable propanol dehydrogenase (EC 1.1.1.-) [imported] - Salmonella enterica subsp. ent  
C:Species: Salmonella enterica subsp. enterica serovar Typh  
A:Note: this species has also been called Salmonella typhi  
C>Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C:Accession: AG0761  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AG0761  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-370 <PAP>  
A:Cross-references: GB:AL513382; PIDN:CAD02413.1; PID:g16503282; GSPDB:GN00176  
C:Genetics:  
A:Gene: pduQ  
C:Superfamily: lactaldehyde reductase; lactaldehyde reductase homology  
C:Keywords: oxidoreductase

Query Match 64.3%; Score 36; DB 2; Length 370;  
Best Local Similarity 63.6%; Pred. No. 96;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALVQMEQLRQ 12  
Db 305 ALVQMEQLRQ 315

## RESULT 21

H83880  
hypothetical protein BH1848 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: H83880  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: AB3650; MUID:20512582; PMID:11058132  
A:Accession: H83880  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-498 <STO>  
A:Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA00567.1; GSPDB:GN00  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH1848

Query Match 64.3%; Score 36; DB 2; Length 498;  
Best Local Similarity 70.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVQMEQL 10  
Db 130 KALVQMEQL 139

## RESULT 22

AS4817  
ATPase ScII, chromosomal scaffold - chicken  
C:Species: Gallus gallus (chicken)  
C>Date: 12-Apr-1995 #sequence\_revision 12-Apr-1995 #text\_change 02-Feb-2001  
C:Accession: AS4817  
R:Salton, N.; Goldberg, I.G.; Wood, E.R.; Earnshaw, W.C.  
J. Cell Biol. 127, 303-318, 1994  
A>Title: ScII: an abundant chromosomal scaffold protein is a member of a family of putative  
A:Reference number: AS4817; MUID:95014725; PMID:7929577  
A:Accession: AS4817  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1189 <SAI>  
A:Cross-references: GB:X80792; NID:g572691; PIDN:CAA56767.1; PID:g572692  
C:Superfamily: chromosome segregation protein SMC1  
C:Keywords: nucleotide binding; P-loop  
P:32-39/Region: nucleotide-binding motif A (P-loop)

Query Match 64.3%; Score 36; DB 2; Length 1189;  
Best Local Similarity 58.3%; Pred. No. 3.1e+02;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KALVQMEQLRQ 12  
Db 820 KALVQMEQLRQ 831

## RESULT 23

T13613  
hypothetical protein 8D8.2 - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
C:Accession: T13613  
R:Papadimitrakis, G.; Spanos, L.; Cox, S.; Sidenkiamos, I.; Louis, C.  
submitted to the EMBL Data Library, April 1999  
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.  
A:Reference number: Z17695  
A:Accession: T13613  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1279 <PAP>  
A:Cross-references: EMBL:AL022018; NID:e1273253; PID:e1426350; PIDN:CAA17685.2  
C:Genetics:  
A:Cross-references: FLYBase:FBgn0024367  
A:Introns: 65/2  
A:Note: EG:8D8.2

Query Match 64.3%; Score 36; DB 2; Length 1279;  
Best Local Similarity 77.8%; Pred. No. 3.4e+02;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VQQLLEQLRQ 12  
Db 643 VQQLLEQLRQ 651

## RESULT 24

T14867  
interactin - slime mold (Dictyostelium discoideum)  
C:Species: Dictyostelium discoideum  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 11-May-2000  
C:Accession: T14867  
R:Rivero, F.J.; Kuspa, A.; Brokamp, R.; Matzner, M.; Noegel, A.A.  
J. Cell Biol. 142, 735-750, 1998  
A>Title: Interactin, an actin-binding protein of the alpha-actinin superfamily in Dictyo  
te.  
A:Reference number: Z18240; MUID:98365468; PMID:9700162

A:Accession: T14867  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1738 <RIV>  
A:Cross-references: EMBL:AF057019; NID:g3549260; PID:g3549261; PIDN:AAC34582.1  
C:Genetics:  
A:Gene: abpd  
A:introns: 173/2; 1680/1

Query Match 64.3%; Score 36; DB 2; Length 1738;  
Best Local Similarity 60.0%; Pred. No. 4.6e+02;  
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 3 KALVQOMEQLRQ 12  
|:|:|:|:|:|:  
Db 1424 LIQEIQLKQ 1433

RESULT 25  
A54103  
centrosome autoantigen PCM-1 - human  
C:Species: Homo sapiens (man)  
C>Date: 02-Aug-1994 #sequence\_revision 02-Aug-1994 #text\_change 08-Oct-1999  
C:Accession: A54103  
R:Balczon, R.; Bao, L.; Zimmer, W.E.  
J. Cell Biol. 124, 783-793, 1994  
A:Title: PCM-1, a 228-kD centrosome autoantigen with a distinct cell cycle distribution.  
A:Reference number: A54103; MUID:94165144; PMID:8120099  
A:Accession: A54103  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2024 <BAL>  
A:Cross-references: GB:L27841; NID:g450276; PIDN:AAA60120.1; PID:g450277

Query Match 64.3%; Score 36; DB 2; Length 2024;  
Best Local Similarity 63.6%; Pred. No. 5.4e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KALVQOMEQLR 11  
|:|:|:|:|:|:  
Db 284 KRMLOQDEQLR 294

RESULT 26  
AE0012  
conserved hypothetical protein YPO0093 [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
C:Accession: AE0012  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titchall, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AE0012  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-79 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC88959.1; PID:g15978204; GSPDB:GN00175  
C:Genetics:  
A:Gene: YPO0093

Query Match 62.5%; Score 35; DB 2; Length 79;  
Best Local Similarity 58.3%; Pred. No. 30;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KALVQOMEQLRQ 12  
|:|:|:|:|:|:  
Db 49 EALVRENQLKQ 60

RESULT 27

H71245  
hypothetical protein PH0223 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 09-Dec-2002  
C:Accession: H71245  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
M.; Ohtuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A:Reference number: A71000; MUID:98344137; PMID:9679194  
A:Accession: H71245  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-190 <KAW>  
A:Cross-references: GB:AP000001; NID:g3236128; PIDN:BAH29295.1; PID:g3256612  
A:Experimental source: strain OT3  
A>Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Gene: PH0223  
C:Superfamily: uncharacterized conserved protein

Query Match 62.5%; Score 35; DB 2; Length 190;  
Best Local Similarity 50.0%; Pred. No. 72;  
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KALVQOMEQLRQ 12  
|:|:|:|:|:|:  
Db 170 KALVEIEIKRK 181

RESULT 28  
S73441  
MG032 homolog B01.orf38 - Mycoplasma pneumoniae (strain ATCC 29342)  
C:Species: Mycoplasma pneumoniae  
A:Variety: ATCC 29342  
C>Date: 26-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 07-Dec-1999  
C:Accession: S73441  
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pickl, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A:Reference number: S73327; MUID:97105885; PMID:8948633  
A:Accession: S73441  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-338 <HIM>  
A:Cross-references: EMBL:AE000014; GB:U00089; NID:g1673770; PIDN:ABD95763.1; PID:g167377  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C:Genetics:  
A:Genetic code: SGC3

Query Match 62.5%; Score 35; DB 2; Length 338;  
Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KALVQOMEQLRQ 12  
|:|:|:|:|:|:  
Db 59 QSLMEQIBQLQ 70

RESULT 29  
E64028  
hypothetical protein H11410 - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 04-Mar-2000  
C:Accession: E64028  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: E64028

A:Statue: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-394 <TIGR>  
A:Cross-references: GB:U32820; GB:U42023; NID:g1574231; PIDN:AAC33058.1; PID:g1574247; I  
C:Genetics:  
A:Start codon: GTG  
C:Superfamily: Haemophilus influenzae hypothetical protein H11410

Query Match 62.5%; Score 35; DB 2; Length 394;  
Best Local Similarity 50.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KALVQMEQLRQ 12  
Db 133 KELMEDMEQMR 144

## RESULT 30

arginosuccinate lyase (EC 4.3.2.1) - Synecocystis sp. (strain PCC 6803)  
N:Alternate names: L-argininosuccinate lyase; protein slr1133  
C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000

A:Accession: S77437  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis  
sp.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S77437

A:Statue: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-461 <KAN>

A:Cross-references: EMBL:D90905; GB:AB001339; NID:g1652360; PIDN:BA17284.1; PID:g165236

A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: argH

C:Superfamily: arginosuccinate lyase

C:Keywords: amidine-lyase; carbon-nitrogen lyase

Query Match 62.5%; Score 35; DB 2; Length 461;  
Best Local Similarity 70.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVQOMEQLRQ 12  
Db 60 LVQGLEQLRQ 69

## RESULT 31

AB04444  
hypothetical protein At2g03070 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: AB04444

R:Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.

Evans, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-769, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB04420; MUID:20083487; PMID:10617197

A:Accession: AB04444

A:Statue: Preliminary

A:Molecule type: DNA

A:Residues: 1-531 <STO>

A:Cross-references: GB:AB002093; NID:g3461831; PIDN:AAC32925.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g03070

A:Map position: 2

Query Match 62.5%; Score 35; DB 2; Length 531;

Best Local Similarity 70.0%; Pred. No. 2.1e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LVQOMEQLRQ 12  
Db 331 LAQOQOQLRQ 340

## RESULT 32

G89813  
DNA polymerase III gamma and tau subunits [imported] - Staphylococcus aureus (strain N31

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C:Accession: G89813

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cu, L.; Oguci

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramoto, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: G89813

A:Statue: Preliminary

A:Molecule type: DNA

A:Residues: 1-565 <KUR>

A:Cross-references: GB:BA000018; PID:g13700368; PIDN:BA041666.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Superfamily: DNA-directed DNA polymerase III gamma chain

Query Match 62.5%; Score 35; DB 2; Length 565;  
Best Local Similarity 70.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVQOMEQLRQ 12  
Db 375 LVQMEQLRQ 384

## RESULT 33

T40124  
kinesin-like motor protein SPBC2D10.21c - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 20-Oct-2000 #text\_change 20-Oct-2000

C:Accession: T40124; T39477

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Taylor, K.; Harris, D.

submitted to the EMBL Data Library, September 1998

A:Reference number: Z21906

A:Accession: T40124

A:Molecule type: DNA

A:Residues: 173-633 <MOO>

A:Cross-references: EMBL:AL031788; PIDN:CAA21179.1; GSPDB:GN00067; SPDB:SPBC2D10.21c

A:Experimental source: strain 972h-; cosmid c2D10

R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lucas, M.; Gallard, C.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z21858

A:Accession: T39477

A:Molecule type: DNA

A:Residues: 1-297 <LVN>

A:Cross-references: EMBL:AL031349; PIDN:CAA20476.1; GSPDB:GN00067; SPDB:SPBC15D4.01c

A:Experimental source: strain 972h-; cosmid c15D4

C:Genetics:

A:Gene: SPDB:SPBC2D10.21c; SPDB:SPBC15D4.01c

A:Map position: 2

Query Match 62.5%; Score 35; DB 2; Length 633;  
Best Local Similarity 58.3%; Pred. No. 2.5e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KALVQMEQLRQ 12  
Db 421 KALTSHLEQLRQ 432



RESULT 34  
E71937  
A:Species: biosynthesis protein - Helicobacter pylori (strain J99)  
C:Species: Helicobacter pylori  
A:Variety: strain J99  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 24-Sep-1999  
C:Accession: E71937  
R:Alt: R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Miller, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
A:Reference number: A71800; MUID:99120557; PMID:9923682  
A:Accession: E71937  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-733 <TOM>  
A:Cross-references: GB:AE001473; GB:AE001439; NID:94154910; PIDN:AD05964.1; PID:9415491  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: flhA  
C:Superfamily: regulatory protein Icrd

Query Match 62.5%; Score 35; DB 2; Length 733;  
Best Local Similarity 72.7%; Pred. No. 2.9e+02;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KALVQMEQLR 11  
Db 696 KALSNQMEQAR 706

RESULT 35  
A64650  
A:Species: biosynthesis protein - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 24-Sep-1999  
C:Accession: A64650  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: A64650  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-733 <TOM>  
A:Cross-references: GB:AE000612; GB:AE000511; NID:92314186; PIDN:AD08087.1; PID:9231418  
C:Superfamily: regulatory protein Icrd

Query Match 62.5%; Score 35; DB 2; Length 733;  
Best Local Similarity 72.7%; Pred. No. 2.9e+02;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KALVQMEQLR 11  
Db 696 KALSNQMEQAR 706

RESULT 36  
C96829  
A:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: C96829  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federjpiel, N.A.; Kaul, S.; White, O.; Alonso, Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizart, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Matti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.M.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: C96829  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-918 <STO>  
A:Cross-references: GB:AE005173; NID:96453864; PIDN:AAF09048.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F19K16.21  
A:Map position: 1

Query Match 62.5%; Score 35; DB 2; Length 918;  
Best Local Similarity 70.0%; Pred. No. 3.6e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 3 LVQMEQLRQ 12  
Db 566 LVQALRELRQ 575

RESULT 37  
AH1906  
A:Species: Nostoc sp. PCC 7120  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AH1906  
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriuchi, Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AH1906  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1043 <RUP>  
A:Cross-references: GB:BA000019; PIDN:BA872759.1; PID:917130147; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all10802

Query Match 62.5%; Score 35; DB 2; Length 1043;  
Best Local Similarity 50.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KALVQMEQLRQ 12  
Db 481 KEIQMESTRQ 492

RESULT 38  
G86301  
A:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: G86301  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federjpiel, N.A.; Kaul, S.; White, O.; Alonso, Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizart, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Matti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, I ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: G86301  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-1413 <STO>  
 A:Cross-references: GB:AE005172; NID:G9989054; PIDN:AGI0817.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1  
 C:Superfamily: retrovirus-related polypeptide

Query Match 62.5%; Score 35; DB 2; Length 1413;  
 Best Local Similarity 70.0%; Pred. No. 5.6e+02;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 LVQOMEQLRQ 12  
 :|||:|  
 Db 164 LVQQLQLRQ 173

## RESULT 39

T31067  
 BIR repeat containing ubiquitin-conjugating enzyme BRUCE - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 21-Jul-2000  
 C:Accession: T31067  
 R:Hauser, H.P.; Bardroff, M.; Pyrowolakis, G.; Jentsch, S.  
 J Cell Biol. 141, 1415-1422, 1998  
 A>Title: A giant ubiquitin-conjugating enzyme related to IAP apoptosis inhibitors.  
 A:Reference number: 220963; PMID:9628897  
 A:Accession: T31067  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-4845 <HNU>  
 A:Cross-references: EMBL:Y17267; NID:G3319989; PIDN:CAA76720.1; PID:G3319990  
 A>Note: localized to the Golgi compartment and the vesicular system  
 C:Keywords: membrane-associated protein

Query Match 62.5%; Score 35; DB 2; Length 4845;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+03;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KALVQOMEQ 9  
 :|||:|  
 Db 2236 KALVEQMEK 2244

## RESULT 40

AF0088  
 hypochetrical protein YPO0718 [imported] - Yersinia pestis (strain CO92)  
 C:Species: Yersinia pestis  
 C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
 C:Accession: AF0088  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
 Nature 413, 523-527, 2001  
 A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
 A:Reference number: AB0001; MUID:21470413; PMID:11586360  
 A:Accession: AF0088  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-124 <KUR>  
 A:Cross-references: GB:AL590842; PIDN:CAC89569.1; PID:G15978802; GSPDB:GN00175  
 C:Genetics:  
 A:Gene: YPO0718

Query Match 60.7%; Score 34; DB 2; Length 124;  
 Best Local Similarity 66.7%; Pred. No. 70;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 4 VQOMEQLRQ 12  
 :|||:|  
 Db 5 LQQLQLRQ 13

Search completed: September 8, 2003, 16:06:44  
 Job time : 21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2003, 16:00:02; Search time 11 Seconds

(without alignments)  
51.302 Million cell updates/sec

Title: US-09-991-809-1

Perfect score: 56  
Sequence: 1 KALVQOMEQLRQ 12

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	100.0	396	1 APA4_HUMAN	P06727 homo sapien
2	56	100.0	401	1 APA4_PAPAN	Q28758 papio anubi
3	56	100.0	429	1 APA4_MACPA	P33621 macaca fasc
4	49	87.5	395	1 APA4_MOUSE	P06728 mus musculi
5	47	83.9	382	1 APA4_PIG	O46409 sus scrofa
6	43	76.8	391	1 APA4_RAT	P02651 rattus norv
7	39	69.6	323	1 MAFB_HUMAN	Q095q3 homo sapien
8	39	69.6	323	1 MAFB_MOUSE	P54841 mus musculi
9	39	69.6	323	1 MAFB_RAT	P54842 rattus norv
10	38	67.9	219	1 FLIZ_BACGU	P35532 bacillus su
11	38	67.9	893	1 YW92_CAEEL	P34531 caenorhabdi
12	37	66.1	788	1 PCAP_HUMAN	Q096tns homo sapien
13	36	64.3	1189	1 SMC2_CHICK	Q09098 gallus gall
14	36	64.3	2090	1 HFC1_MESAU	P51611 mesocricetu
15	35	62.5	338	1 YA39_MYCPN	P75075 mycoplasma
16	35	62.5	394	1 YE10_HABIN	P44184 haemophilus
17	35	62.5	447	1 PURA_SYNEL	O681g2 synechococ
18	35	62.5	461	1 ARLY_SYNT3	P73257 synechocyst
19	35	62.5	463	1 ARLY_SYNEL	Q081w0 synechococ
20	35	62.5	733	1 FLHA_HELPJ	O92m40 helicobacte
21	35	62.5	733	1 FLHA_HELPJ	O06758 helicobacte
22	35	62.5	956	1 KP5C_MOUSE	P28738 mus musculi
23	35	62.5	957	1 KP5C_MOUSE	O60282 homo sapien
24	35	62.5	1938	1 MYHD_HUMAN	Q90ux3 homo sapien
25	35	62.5	4829	1 BIR6_HUMAN	Q9nrt09 homo sapien
26	34	60.7	213	1 YBGJ_HABIN	P44299 haemophilus
27	34	60.7	264	1 CBIX_SALTY	Q05592 salmoneilla
28	34	60.7	473	1 SECA_BACFI	P96313 bacillus fi
29	34	60.7	492	1 SE81_MOUSE	P58006 mus musculi
30	34	60.7	514	1 SAD1_SCHRO	Q09825 schizosach
31	34	60.7	541	1 PRPR_SALTY	P78839 salmoneilla
32	34	60.7	569	1 APB3_RAT	O70248 rattus norv
33	34	60.7	708	1 CA02_CANTR	P06598 candida tro

34	34	60.7	796	1 DECI1_YEAST	Q12387 saccharomyc
35	34	60.7	861	1 GCR3_YEAST	P34160 saccharomyc
36	34	60.7	1085	1 IFH1_YEAST	P33920 saccharomyc
37	33	58.9	244	1 AG13_ARATH	Q38837 arabidopsis
38	33	58.9	269	1 TRPA_BACST	P19867 bacillus st
39	33	58.9	443	1 SR54_PYRU	Q8u070 pyrococcus
40	33	58.9	448	1 RMUC_ZYMO	Q9reg3 zymomonas m
41	33	58.9	459	1 CCMH_HABIN	P46458 haemophilus
42	33	58.9	459	1 EX7L_YERPE	O82cu2 yersinia pe
43	33	58.9	501	1 MYSU_RABIT	O99105 oryctolagus
44	33	58.9	544	1 UBIB_VIBRU	O8ddq1 vibrio vuln
45	33	58.9	824	1 CC27_HUMAN	P30260 homo sapien

## ALIGNMENTS

RESULT 1  
ID APA4\_HUMAN STANDARD; PRT; 396 AA.  
AC P06727;  
DT 01-JAN-1988 (Rel. 05, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Apolipoprotein A-IV precursor (Apo-AIV).  
GN APOA4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Intestine;  
RX MEDLINE=89194198; PubMed=2930771;  
RA Yang C., Gu Z.W., Xiong W., Rosseneu M., Yang H.X., Lee B.M.,  
Gocto A.M. Jr., Chan L.;  
RT "The primary structure of human apolipoprotein A-IV";  
RL Biochim. Biophys. Acta 1002:231-237(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87041474; PubMed=3095836;  
RA Karathanasis S.K., Oetgen P., Haddad I.A., Antonarakis S.E.;  
RT "Structure, evolution, and polymorphisms of the human apolipoprotein  
A4 gene (APOA4)";  
RL Proc. Natl. Acad. Sci. U.S.A. 83:8457-8461(1986).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86296629; PubMed=3755616;  
RA Karathanasis S.K., Yunis I.;  
RT "Structure, evolution, and tissue-specific synthesis of human  
apolipoprotein AIV";  
RL Biochemistry 25:3962-3970(1986).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87250378; PubMed=3036793;  
RA Elshourbagy N.A., Walker D.W., Paik Y.K., Boguski M.S., Freeman M.,  
Gordon J.I., Taylor J.W.;  
RT "Structure and expression of the human apolipoprotein A-IV gene";  
RL J. Biol. Chem. 262:7973-7981(1987).  
RN [5]  
RP SEQUENCE OF 21-396 FROM N.A.  
RX MEDLINE=86111885; PubMed=3080432;  
RA Elshourbagy N.A., Walker D.W., Boguski M.S., Gordon J.I., Taylor J.M.;  
RT "The nucleotide and derived amino acid sequence of human  
apolipoprotein A-IV mRNA and the close linkage of its gene to the  
genes of apolipoproteins A-I and C-III";  
RL J. Biol. Chem. 261:1998-2002(1986).  
RN [6]  
RP SIGNAL SEQUENCE CLEAVAGE SITE  
RX MEDLINE=84161950; PubMed=6706947;  
RA Gordon J.I., Bisgater C.L., Sims H.F., Sachdev O.P., Glickman R.M.,  
Strauss A.W.;  
RT "Biosynthesis of human preapolipoprotein A-IV";

RL J. Biol. Chem. 259:468-474(1984).  
 RN [7].  
 RN REVIEW ON POLYMORPHISM.  
 RA Lohse P., Brewer H.B. Jr.;  
 RT "genetic polymorphism of apolipoprotein A-IV";  
 RL Curr. Opin. Lipidol. 2:90-95(1991).  
 RN [8].  
 RP VARIANT A-IV\*2.  
 RX MEDLINE=90277616; PubMed=2351649;  
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RT "genetic polymorphism of human plasma apolipoprotein A-IV is due to  
 RT nucleotide substitutions in the apolipoprotein A-IV gene.";  
 RL J. Biol. Chem. 265:10061-10064(1990).  
 RN [9].  
 RP VARIANTS A-IV\*0 AND A-IV\*3.  
 RX MEDLINE=90324273; PubMed=1973689;  
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RT "Three genetic variants of human plasma apolipoprotein A-IV: apoA-IV-  
 RT 1(Thr-347-->Ser), apoA-IV-0(Lys-167-->Glu,Gln-360-->His), and apoA-IV-  
 RT 3(Glu-165-->Lys).";  
 RL J. Biol. Chem. 266:13513-13518(1991).  
 RN [11].  
 RP ERRATUM.  
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RL J. Biol. Chem. 266:19866-19866(1991).  
 RN [12].  
 RP VARIANT MET-13.  
 RX MEDLINE=92238494; PubMed=1359197;  
 RA von Eckardstein A., Funke H., Schulte M., Erren M., Schulte H.,  
 RA Assmann G.;  
 RT "Nonsynonymous polymorphic sites in the apolipoprotein (apo) A-IV  
 RT gene are associated with changes in the concentration of apo B- and  
 RT apo A-I-containing lipoproteins in a normal population.";  
 RL Am. J. Hum. Genet. 50:1115-1128(1992).  
 RN [13].  
 RP VARIANT SER-147.  
 RX MEDLINE=92144647; PubMed=1737067;  
 RA Tenkansen H., Koskinen P., Metsu J., Baumann M., Lukka M.,  
 RA Kauppinen-Makela R., Konttinen K., Taskiran M.R., Manttari M.,  
 RA Manninen V., Ehnholm C.;  
 RT "A novel polymorphism of apolipoprotein A-IV is the result of an  
 RT asparagine to serine substitution at residue 127.";  
 RL Biochim. Biophys. Acta 1138:27-33(1992).  
 RN [14].  
 RP VARIANT A-IV\*5.  
 RX MEDLINE=93138374; PubMed=1487136;  
 RA Kamboh M.I., Williams E.R., Law J.C., Aston C.E., Bunker C.H.,  
 RA Ferrell R.E., Pollitzer W.S.;  
 RT "Molecular basis of a unique African variant (A-IV 5) of human  
 RT apolipoprotein A-IV and its significance in lipid metabolism.";  
 RL Genet. Epidemiol. 9:379-388(1992).  
 RN [15].  
 RP VARIANTS BUNAPEST-2 LYS-64; BUNAPEST-1 CYS-305 AND SER-367.  
 RX MEDLINE=95245341; PubMed=7728150;  
 RA Menzel H.-J., Dieplinger H., Sanderholzer C., Karadi I., Utermann G.,  
 RA Csaszar A.;  
 RT "Apolipoprotein A-IV polymorphism in the Hungarian population: gene  
 RT frequencies, effect on lipid levels, and sequence of two new  
 RT variants.";  
 RL Hum. Mutat. 5:58-65(1995).  
 RN [16].  
 RP VARIANTS SEATTLE-3 SER-161; SEATTLE-1 LEU-178 AND SEATTLE-2 GLN-264,  
 RX MEDLINE=97114287; PubMed=8956036;  
 RA Deeb S.S., Nevil D.N., Iwasaki L., Brunzell J.D.;  
 RT "Two novel apolipoprotein A-IV variants in individuals with familial  
 RT combined hyperlipidemia and diminished levels of lipoprotein lipase

```

RT activity." ;
RL Hum. Mutat. 8:319-325 (1996) .
RN [17]
RP VARIANT HIS-380.
RX MEDLINE=99318094; PubMed=10391210;
RA Halushka M.K., Fan J.-B., Bentley K., Hsie L., Shen N., Weder A.,
RA Cooper R., Lipshutz R., Chakravarti A. ;
RT "Patterns of single-nucleotide polymorphisms in candidate genes for
RT blood-pressure homeostasis." ;
RL Nat. Genet. 22:239-247(1999) .
CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN
CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR
CC COMPONENT OF HDL AND CHYLOMICRONS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE INTESTINE.
CC SECRETED IN PLASMA.
CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH
CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-
CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL
CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.
CC -1- POLYMORPHISM: EIGHT ALLELES HAVE BEEN CHARACTERIZED (APOA-IV*0
CC TO APOA-IV*7). APOA-IV*1 IS THE MAJOR ALLELE (90%), IV*2 IS ALSO
CC COMMON (8%). THE OTHERS ARE RARE ALLELES.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk) .
CC -----
DR EMBL; M14642; AAA51745.1; -.
DR EMBL; X13629; CAAS1955.1; -.
DR EMBL; M14566; AAA51748.1; -.
DR EMBL; J02758; AAA96731.1; -.
DR EMBL; M13654; AAA51744.1; -.
DR PIR; A94137; LPHU44.
DR HSSP; P02649; LEE4.
DR SMSS-2DPAGE; P06727; HUMAN.
DR GeneW; HGNC:602; APOA4.
DR MIM; 107690; -.
DR GO; GO:0005320; F:apolipoprotein; TAS.
DR GO; GO:0008015; P:circulation; TAS.
DR GO; GO:0006869; P:lipid transport; TAS.
DR InterPro; IPR00074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 2.
KW Plasma; Lipid transport; HDL; Chylomicron; Repeat; Signal;
KW Polymorphism.
FT SIGNAL 1 20
FT CHAIN 1 396
FT DOMAIN 33 330
FT REPEAT 33 54
FT REPEAT 60 81
FT REPEAT 82 103
FT REPEAT 115 136
FT REPEAT 137 158
FT REPEAT 159 180
FT REPEAT 181 202
FT REPEAT 203 224
FT REPEAT 225 246
FT REPEAT 247 268
FT REPEAT 269 286
FT REPEAT 287 308
FT REPEAT 309 330
FT DOMAIN 372 389
FT VARIANT 13 13
FT VARIANT 44 44
FT VARIANT 44 44
E -> K (IN BUDAPEST-2) .
EFTID=VAR_000626

```

Query Match 100.0%; Score 56; DB 1; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.014;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KALVOQMEOLRQ 12  
 |||||  
 DB 316 KALVOQMEOLRQ 327

## RESULT 2

AP04\_PAPAN STANDARD; PRT; 401 AA.  
 AC Q28758;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Apolipoprotein A-IV precursor (Apo-AIV) (Fragment).  
 GN APOA4.  
 OS Papio anubis (Olive baboon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Papio.  
 OX NCBI\_TaxID=9555;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Intestine;  
 RX MEDLINE=93340170; PubMed=8101842;  
 RA Hixson J.E., Kammerer C.M., Mott G.E., Britten M.L., Birnbaum S.,  
 RA Powers P.K., Vandeberg J.L.;  
 RT "Baboon apolipoprotein A-IV. Identification of Lys-76-->Glu that  
 RT distinguishes two common isoforms and detection of length  
 RT polymorphisms at the carboxyl terminus".  
 RL J. Biol. Chem. 268:15667-15673(1993).  
 CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND  
 CC CARBOHYDRATE METABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN  
 CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR  
 CC COMPONENT OF HDL AND CHYLOMICRONS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE INTESTINE.  
 CC SECRETED IN PLASMA.  
 CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH  
 CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-  
 CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-  
 CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY  
 CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECTITHIN:CHOLESTEROL  
 CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.  
 CC -1- POLYMORPHISM: THERE ARE TWO COMMON APOA-IV ISOFORMS, I (SHOWN  
 CC HERE) AND E. THE I ISOFORM IS ASSOCIATED WITH HIGHER LEVELS OF  
 CC HIGH DENSITY LIPOPROTEIN-C ON A HIGH CHOLESTEROL, SATURATED FAT  
 CC DIET.  
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; L13174; AAA5379.1; -;  
 DR HSSP; P02649; INFN.  
 DR InterPro; IPR000074; Apolipoprotein.  
 DR Pfam; PF01442; Apolipoprotein; 2.  
 KW Plasma; Lipid transport; HDL; Chylomicron; Repeat; Signal;  
 KW Polymorphism.  
 FT NON TER 1  
 FT SIGNAL 4  
 FT CHAIN 5 401  
 FT DOMAIN 17 314  
 FT REPEAT 17 38  
 FT REPEAT 44 65  
 FT REPEAT 2;

FT REPEAT 66 87 3  
 FT REPEAT 99 120 4  
 FT REPEAT 121 142 5  
 FT REPEAT 143 164 6  
 FT REPEAT 165 186 7  
 FT REPEAT 187 208 8  
 FT REPEAT 209 230 9  
 FT REPEAT 231 252 10  
 FT REPEAT 253 270 11  
 FT REPEAT 271 292 12  
 FT REPEAT 293 314 13  
 FT DOMAIN 356 394  
 FT VARIANT 80  
 SQ SEQUENCE 401 AA; 46538 MW; 0A76D1284AA9837F CRC64;  
 K->E (IN ISOFORM E).

Query Match 100.0%; Score 56; DB 1; Length 401;  
 Best Local Similarity 100.0%; Pred. NO. 0.015;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KALVOQMEOLRQ 12  
 |||||  
 DB 300 KALVOQMEOLRQ 311

## RESULT 3

AP04\_MACFA STANDARD; PRT; 429 AA.  
 AC P33621;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Apolipoprotein A-IV precursor (Apo-AIV).  
 GN APOA4.  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leukocyte;  
 RX MEDLINE=93192330; PubMed=8448212;  
 RA Osada J., Pocovi M., Nicolosi R.J., Schaefer E.J., Ordovas J.M.;  
 RT "Nucleotide sequences of the Macaca fascicularis apolipoprotein C-II  
 RT and A-IV genes".  
 RL Biochim. Biophys. Acta 1172:335-339(1993).  
 CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND  
 CC CARBOHYDRATE METABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN  
 CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR  
 CC COMPONENT OF HDL AND CHYLOMICRONS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Secreted in plasma.  
 CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH  
 CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-  
 CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-  
 CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY  
 CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECTITHIN:CHOLESTEROL  
 CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.  
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X68361; CAA48421.1; -;  
 DR PIR; S30195; S29565.  
 DR HSSP; P02649; INFN.  
 DR InterPro; IPR000074; Apolipoprotein.  
 DR Pfam; PF01442; Apolipoprotein; 2.

Query Match	Best Local Similarity	Score 56;	DB 1;	Length 429;	
Matches 12;	Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0
1 KALVQOMEQLRQ 12	100.0%;	56;	DB 1;	Length 429;	
316 KALVQOMEQLRQ 327	100.0%;	Pred. No. 0.016;			
SEQUENCE	429 AA;	49876 MW;	3D458F51D0DB60C	CRC64;	
	372	420	GUU/GUN-RICH.		
FT SIGNAL	1	20			
FT CHAIN	21	429			
FT DOMAIN	33	330			
FT REPEAT	33	54			
FT REPEAT	60	81			
FT REPEAT	82	103			
FT REPEAT	115	136			
FT REPEAT	137	158			
FT REPEAT	159	180			
FT REPEAT	181	202			
FT REPEAT	203	224			
FT REPEAT	225	246			
FT REPEAT	247	268			
FT REPEAT	269	286			
FT REPEAT	287	308			
FT REPEAT	309	330			
FT REPEAT	372	420			
FT DOMAIN	429 AA;	49876 MW;	3D458F51D0DB60C	CRC64;	
FT SIGNAL	1	20			
FT CHAIN	21	429			
FT DOMAIN	33	330			
FT REPEAT	33	54			
FT REPEAT	60	81			
FT REPEAT	82	103			
FT REPEAT	115	136			
FT REPEAT	137	158			
FT REPEAT	159	180			
FT REPEAT	181	202			
FT REPEAT	203	224			
FT REPEAT	225	246			
FT REPEAT	247	268			
FT REPEAT	269	286			
FT REPEAT	287	308			
FT REPEAT	309	330			
FT REPEAT	372	420			
FT DOMAIN	429 AA;	49876 MW;	3D458F51D0DB60C	CRC64;	

```

CC CC ACYLTRANSFERASE (LOCAL) ACTIVATING ABILITIES.
CC CC -1- POLYMORPHISM: THERE IS A POLYMORPHISM WITHIN A SERIES OF IMPERFECT
CC CC REPEATS ENCODING THE SEQUENCE E-O-(AV)-O. INSERTIONS OR DELETIONS
CC CC OF 12 NUCLEOTIDES HAVE GIVEN RISE TO THREE FORMS CHARACTERIZED BY
CC CC THREE (129), FOUR (C57BL/6), OR FIVE (M.CASTANEUS) COPIES OF THE
CC CC REPEAT UNIT.
CC CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb.ch/announce/
CC CC or send an email to license@isb.ch).
CC CC -----
CC CC EMBL; M13966; AAA37253.1; -.
CC CC EMBL; M64249; AAA37215.1; -.
CC CC EMBL; M64248; AAA37214.1; -.
CC CC PIR; A25281; A25281.
CC CC PIR; A40892; A40892.
CC CC PIR; B40892; B40892.
CC CC MGD; MGI:88051; APOA4.
CC CC DR GO; GO:0030300; P:regulation of cholesterol absorption; IMP.
CC CC InterPro; IPR000074; Apolipoprotein.
CC CC DR Pfam; PF01442; Apolipoprotein_2
CC CC KW Plasma; Lipid transport; HDL; Chylomicron; Repeat; Signal;
CC CC Polymorphism.
CC CC KW SIGNAL 1 20
CC CC FT CHAIN 21 395 APOLOPROTEIN A-IV.
CC CC FT DOMAIN 33 330 13 X 22 AA APPROXIMATE TANDEM REPEATS.
CC CC FT REPEAT 33 54 1.
CC CC FT REPEAT 60 81 2.
CC CC FT REPEAT 82 103 3.
CC CC FT REPEAT 115 136 4.
CC CC FT REPEAT 137 158 5.
CC CC FT REPEAT 159 180 6.
CC CC FT REPEAT 181 202 7.
CC CC FT REPEAT 203 224 8.
CC CC FT REPEAT 225 246 9.
CC CC FT REPEAT 247 268 10.
CC CC FT REPEAT 269 286 11.
CC CC FT REPEAT 287 308 12.
CC CC FT REPEAT 309 330 13.
CC CC FT DOMAIN 366 389 GLU/GLN-RICH.
CC CC FT VARIANT 382 385 MISSING (IN SOME STRAINS).
CC CC FT CONFLICT 15 15 MISSING (IN REF. 1).
CC CC FT CONFLICT 63 63 O -> K (IN REF. 1).
CC CC FT CONFLICT 207 207 E -> R (IN REF. 1).
CC CC FT CONFLICT 288 288 RQ -> A (IN REF. 1).
CC CC FT CONFLICT 294 295 RO -> KA (IN REF. 1).
CC CC FT CONFLICT 315 316 NK -> GG (IN REF. 1).
CC CC SQ SEQUENCE 395 AA; 45029 MW; 5FE27D023626257 CRC64;
CC CC -----
CC CC Query Match 87.5%; Score 49; DB 1; Length 395;
CC CC Best Local Similarity 83.3%; Pred. No. 0.24;
CC CC Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0
CC CC
CC CC QY 1 KALVOOMEQRLQ 12
CC CC |||||:||||
CC CC DB 316 KALVOOLEQFRQ 327
CC CC
CC CC RESULT 5
CC CC APOA4 PIG STANDARD; PRT; 382 AA.
CC CC ID _APO4_PIG
CC CC AC Q46409;
CC CC DT 15-DEC-1998 (Rel. 37, Created)
CC CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
CC CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC CC DE Apolipoprotein A-IV precursor (Apo-AIV).
CC CC GN APOA4.
CC CC OS Sus scrofa (Pig).

```

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 CC NCBI\_TaxID=9823;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC Osada J., Iturza-Gadea M., Calleja L., Gonzalez N., Pineiro A.,  
 CC RL FUNCTION (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Secreted in plasma.  
 CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH  
 CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-  
 CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-  
 CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY  
 CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL  
 CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.  
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AJ222966; CAA11020.1; -;  
 CC DR HSSP; P32851; 1BR0.  
 CC DR InterPro; IPR000074; Apolipoprotein.  
 CC Pfam; PF01442; Apolipoprotein; 2.  
 CC KW Plasma; Lipid transport; HDL; Chylomicron; Repeat; Signal.  
 CC FT SIGNAL 1 20  
 CC FT CHAIN 1 382 APOLIPOPROTEIN A-IV.  
 CC FT DOMAIN 33 330 13 X 22 AA APPROXIMATE TANDEM REPEATS.  
 CC FT REPEAT 33 54 1.  
 CC FT REPEAT 60 81 2.  
 CC FT REPEAT 82 103 3.  
 CC FT REPEAT 115 136 4.  
 CC FT REPEAT 137 158 5.  
 CC FT REPEAT 159 180 6.  
 CC FT REPEAT 181 202 7.  
 CC FT REPEAT 203 224 8.  
 CC FT REPEAT 225 246 9.  
 CC FT REPEAT 247 268 10.  
 CC FT REPEAT 269 286 11.  
 CC FT REPEAT 287 308 12.  
 CC FT REPEAT 309 330 13.  
 CC FT DOMAIN 360 377 GLU/GLN-RICH.  
 CC SO SEQUENCE 382 AA; 43294 MW; 2ACAA88A024379EF CRC64;  
 CC  
 CC Query Match 83.9%; Score 47; DB 1; Length 382;  
 CC Best Local Similarity 83.3%; Pred. NO. 0.51;  
 CC Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC Oy 1 KALVOOMOLRO 12  
 CC Db 316 KALVOOVEDLRO 327  
 CC  
 CC RESULT 6  
 CC APOA4\_RAT STANDARD; PRT; 391 AA.  
 CC AC P02651;  
 CC DT 21-JUL-1986 (Rel. 01, Created)  
 CC DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Apolipoprotein A-IV precursor (Apo-AIV).  
 CC GN APOA4.  
 CC OS Rattus norvegicus (Rat).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 CC NCBI\_TaxID=10116;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC MEDLINE=86196059; PubMed=3009456;  
 CC RA Boguski M.S., Birkenmeier E.H., Eshourbagy N.A., Taylor J.M.,  
 CC RA Gordon J.I.;  
 CC "Evolution of the apolipoproteins. Structure of the rat apo-A-IV gene  
 CC and its relationship to the human genes for apo-A-I, C-III, and E.";   
 CC J. Biol. Chem. 261:6398-6407(1986).  
 CC [2]  
 CC SEQUENCE FROM N.A.  
 CC MEDLINE=84298074; PubMed=6591177;  
 CC RA Boguski M.S., Eshourbagy N.A., Taylor J.M., Gordon J.I.;  
 CC "Rat apolipoprotein A-IV contains 13 tandem repetitions of a 22-amino  
 CC acid segment with amphipathic helical potential.";   
 CC Proc. Natl. Acad. Sci. U.S.A. 81:5021-5025(1984).  
 CC [3]  
 CC SEQUENCE FROM N.A.  
 CC MEDLINE=87008540; PubMed=3020028;  
 CC RA Haddad I.A., Ordovas J.M., Fitzpatrick T., Karathanasis S.K.;  
 CC "Linkage, evolution, and expression of the rat apolipoprotein A-I, C-  
 CC III, and A-IV genes.";   
 CC J. Biol. Chem. 261:13268-13277(1986).  
 CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND  
 CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN  
 CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR  
 CC COMPONENT OF HDL AND CHYLOMICRONS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Secreted in plasma.  
 CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH  
 CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-  
 CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-  
 CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY  
 CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL  
 CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.  
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M00002; AA85909.1; -;  
 CC DR EMBL; J02588; AAA40747.1; -;  
 CC DR EMBL; M13508; AAA40748.1; -;  
 CC DR PIR; A03095; LPRPA4.  
 CC DR InterPro; IPR000074; Apolipoprotein.  
 CC Pfam; PF01442; Apolipoprotein; 2.  
 CC KW Plasma; Lipid transport; HDL; Chylomicron; Repeat; Signal.  
 CC FT SIGNAL 1 20  
 CC FT CHAIN 1 391 APOLIPOPROTEIN A-IV.  
 CC FT DOMAIN 33 330 13 X 22 AA APPROXIMATE TANDEM REPEATS.  
 CC FT REPEAT 33 54 1.  
 CC FT REPEAT 60 81 2.  
 CC FT REPEAT 82 103 3.  
 CC FT REPEAT 115 136 4.  
 CC FT REPEAT 137 158 5.  
 CC FT REPEAT 159 180 6.  
 CC FT REPEAT 181 202 7.  
 CC FT REPEAT 203 224 8.  
 CC FT REPEAT 225 246 9.  
 CC FT REPEAT 247 268 10.  
 CC FT REPEAT 269 286 11.  
 CC FT REPEAT 287 308 12.  
 CC FT REPEAT 309 330 13.  
 CC FT DOMAIN 374 385 GLU/GLN-RICH.  
 CC FT VARIANT 253 0 -> H.  
 CC SO SEQUENCE 391 AA; 44456 MW; 24095004A809201D CRC64;

```

Query Match          76.8%; Score 43; DB 1; Length 391;
Best Local Similarity 81.8%; Pred. No. 2.6;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 LVQOMEQRLQ 12
Db      317 LVQOMEQRLQ 327

RESULT 7
MAFB_HUMAN          STANDARD; PRT; 323 AA.
ID MAFB_HUMAN
AC Q9Y5Q3; Q9H1F1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcription factor MafB (V-maf musculoaponeurotic fibrosarcoma
DE oncogene homolog B).
GN MAFB OR KRM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=99375320; PubMed=10444328;
RA Wang P.W., Eisenbart J.D., Cordes S.P., Barsh G.S., Stoeffel M.,
RA Le Beau M.M.,
RA "Human KRM1 (MAFB): cDNA cloning, genomic structure, and evaluation as
RT a candidate tumor suppressor gene in myeloid leukemias."
RL Genomics 59:275-281(1999).
[2]
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.B., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Gardner C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cordy N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths M.N.D., Griffiths G., Gillman A., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.F., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehaesialho M.H., Levertha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., Murray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Senra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.,
RT "The DNA sequence and comparative analysis of human chromosome 20."
RL Nature 414:865-871(2001).
[3]
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

```

```

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Plays a pivotal role in regulating lineage-specific
CC hematopoiesis by repressing ETS1-mediated transcription of
CC erythroid-specific genes in myeloid cells (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the bZIP family. Maf subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AF134157; AAD30106.1; -
DR EMBL; AL035665; CAB75863.1; -
DR EMBL; BC028098; AAH28098.1; -
DR EMBL; BC036689; AAH36689.1; -
DR Genew; HGNC:6408; MAFB.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0007423; P:sensory organ development; TAS.
DR InterPro; IPR004827; TF_bZIP.
DR InterPro; IPR004826; TF_Maf.
DR Pfam; PF03131; bZIP_Maf; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS50217; BZIP; 1.
DR PROSITE; PS00036; BZIP_BASIC; FALSE NEG.
KW Transcription regulation; Repressor; DNA-binding; Nuclear protein.
FT DNA_BIND 238
FT DOMAIN 264
FT DOMAIN 266
FT DOMAIN 287
FT DOMAIN 131
FT DOMAIN 143
FT DOMAIN 158
FT DOMAIN 167
FT CONFLICT 52
FT CONFLICT 241
FT CONFLICT 241
SQ SEQUENCE 323 AA; 35792 MW; A0F3C09F836CB16 CRC64;

Query Match          69.6%; Score 39; DB 1; Length 323;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 LVQOMEQRLQ 12
Db      273 LVQOMEQRLQ 282

RESULT 8
MAFB_MOUSE          STANDARD; PRT; 323 AA.
ID MAFB_MOUSE
AC P54841;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcription factor MafB (V-maf musculoaponeurotic fibrosarcoma
DE oncogene homolog B) (Transcription factor MAF1) (Segmentation protein
DE KR) (Kreisel).
GN MAFB OR MAF1 OR KRM1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```



```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95094266; PubMed=8001130;
RA Cordes S.P., Barsh G.S.;
RT "The mouse segmentation gene kr encodes a novel basic domain-leucine
zipper transcription factor.";
RL 79:1025-1034(1994).
CC -1- FUNCTION: Plays a pivotal role in regulating lineage-specific
hematopoiesis by repressing ETS1-mediated transcription of
erythroid-specific genes in myeloid cells.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN KIDNEY, GUT, LUNG, AND BRAIN.
CC -1- DEVELOPMENTAL STAGE: DETECTABLE AT 8.0 DPC (ONE SOMITE) AS A BAND
IN THE CAUDAL HINDBRAIN, AND BY 8.5 DPC (SIX TO EIGHT SOMITES),
THE HIGH LEVEL DOMAIN EXHIBITS A SHARP ROSTRAL EDGE COINCIDENT
WITH THE R4/R5 BOUNDARY AND A DIFFUSE CAUDAL EDGE LOCATED MIDWAY
THROUGH R6.
CC -1- SIMILARITY: Belongs to the bzip family. Maf subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL, L36435; AAA65689.1; -
DR PIR, I49529; I49529.
DR TRANSFAC, T01439; -.
DR MGD, MGI:104555; Mafp.
DR InterPro: IPR004827; TF_bZIP.
DR InterPro: IPR004826; TF_Maf.
DR Pfam, PF03131; bZIP_Maf 1.
DR SMART, SM00338; BRLZ; 1.
DR PROSITE, PS50217; BZIP; 1.
DR PROSITE, PS00036; BZIP_BASIC; FALSE NEG.
DR Transcription regulation; Repressor; DNA-binding; Nuclear protein.
KM DNA BIND 238 264
FT DOMAIN 266 287 LEUCINE-ZIPPER.
FT DOMAIN 131 143 POLY-HIS.
FT DOMAIN 158 167 POLY-HIS.
FT MUTAGEN 248 248 N->S: LOSS OF TRANSCRIPTIONAL ACTIVITY.
SQ SEQUENCE 323 AA; 35809 MW; D77AE07ABD9C2AD2 CRC64;

Query Match 69.6%; Score 39; DB 1; Length 323;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 LVQOMEQRLQ 12
DB 273 LIQVEQLKQ 282

RESULT 9
MAFB_RAT STANDARD; PRT; 323 AA.
AC P54812;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcription factor MafB (V-maf musculoaponeurotic fibrosarcoma
oncogene homolog B) (Transcription factor MafB).
GN MAFB OR MAF1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mietar; TISSUE=Liver;
RX MEDLINE=97190228; PubMed=9038363;
RA Sakai M., Imaki J., Yoshida K., Ogata A., Matsushima-Hibaya Y.,

```

```

RA Kuboki Y., Nishizawa M., Nishi S.;
RT "Rat maf related genes: specific expression in chondrocytes, lens and
RT spinal cord.";
RL Oncogene 14:745-750(1997).
CC -1- FUNCTION: Plays a pivotal role in regulating lineage-specific
hematopoiesis by repressing ETS1-mediated transcription of
erythroid-specific genes in myeloid cells (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the bzip family. Maf subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL, U56241; AAB50062.1; -
DR InterPro: IPR004827; TF_bZIP.
DR InterPro: IPR004826; TF_Maf.
DR Pfam, PF03131; bZIP_Maf 1.
DR SMART, SM00338; BRLZ; 1.
DR PROSITE, PS50217; BZIP; 1.
DR PROSITE, PS00036; BZIP_BASIC; FALSE NEG.
KM Transcription regulation; Repressor; DNA-binding; Nuclear protein.
FT DNA BIND 238 264
FT DOMAIN 266 287 LEUCINE-ZIPPER.
FT DOMAIN 131 143 POLY-HIS.
FT DOMAIN 158 167 POLY-HIS.
FT DOMAIN 194 201 POLY-ALA.
SQ SEQUENCE 323 AA; 35792 MW; 6E386340D1F840A5 CRC64;

Query Match 69.6%; Score 39; DB 1; Length 323;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 LVQOMEQRLQ 12
DB 273 LIQVEQLKQ 282

RESULT 10
FLIZ_BACSU STANDARD; PRT; 219 AA.
AC P35536;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Flagellar biosynthetic protein fliz precursor.
GN FLIZ OR CHEA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / OI1085;
RX MEDLINE=92283757; PubMed=1597417;
RA Bischoff D.S., Weinreich M.D., Ordal G.W.;
RT "Nucleotide sequences of Bacillus subtilis flagellar biosynthetic
RT genes fliz and flizq and identification of a novel flagellar gene,
RT fliz2".
RL J. Bacteriol. 174:4017-4025(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Boloitin A., Borchert S.,
RA Borrielle R., Bouteiller L., Brans A., Braun N., Brigelli S.C., Bron S.,
RA Brouillet S., Bruschel C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devigne K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,

```

RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
RA Giuseppe G., Guy B.J., Haga K., Haelegh J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karmata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kufita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sorokin J., Sekowski A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
RA Sorokin M., Taccetti E., Takagi T., Takahashi H., Takemaru K.,  
RA Takuchi M., Yamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambut R., Wedler E., Wedler H., Weltzenegger T.,  
RA Winters P., Wipac A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,  
RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
subtilis."  
RL Nature 390:249-256(1997).  
CC -1- FUNCTION: MAY BE A STRUCTURAL COMPONENT OF THE FLAGELLUM THAT  
CC ANCHORS THE ROD TO THE MEMBRANE.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
CC EMBL; M87005; AAA22452.1; -  
DR EMBL; Z99112; CAB13507.1; -  
DR PIR; B41886; B41886.  
DR Subtilist; BG10259; FL1Z.  
KM Flagella; Signal; Transmembrane; Complete proteome.  
FT SIGNAL 1 26  
FT CHAIN 27 219  
FT TRANSMEM 71 96  
FT POTENTIAL.  
SQ SEQUENCE 219 AA; 24871 MW; 6FE82AF8B1DAE8E CRC64;  
  
Query Match 67.9%; Score 38; DB 1; Length 219;  
Best Local Similarity 58.3%; Pred. No. 11;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 KALVOCMEQLRQ 12  
Db 194 KALKEQLEELKQ 205  
  
RESULT 11  
YW92 CAEEL STANDB; 893 AA.  
AC P34531; P34532; P34533; PRT; 893 AA.  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical protein M01A8.2 in chromosome III.  
GN M01A8.2  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodexinae; Caenorhabditis.  
OX NCB1\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=bristol N2;  
RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Ainscough R., Anderson K., Baynes C., Berke M.,  
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
RA Johnston L., Jones M., Kerhaw J., Kirren J., Lalasser N.,  
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,  
RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,  
RA Sulten J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
RA Wohldman P.,  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans."  
RL Nature 368:32-38(1994).  
CC -1- SIMILARITY: Contains 1 CAP-Gly domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
CC EMBL; Z27081; CAA81607.1; -  
DR PIR; G88551; G88551.  
DR WormPep; M01A8.2; CE03491.  
DR InterPro; IPR000938; CAP-Gly.  
DR Pfam; PF01302; CAP-Gly; 1.  
DR PROSITE; PS00845; CAP-Gly; 1.  
DR PROSITE; PS50245; CAP-Gly; 2; 1.  
KW Hypothetical protein; Coiled coil.  
FT DOMAIN 39 81  
FT DOMAIN 522 696  
FT DOMAIN 729 756  
FT COILED COIL (POTENTIAL).  
FT COILED COIL (POTENTIAL).  
SQ SEQUENCE 893 AA; 99997 MW; 464F2962B36C8B1 CRC64;  
  
Query Match 67.9%; Score 38; DB 1; Length 893;  
Best Local Similarity 66.7%; Pred. No. 43;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 KALVOCMEQLRQ 12  
Db 725 KILVQIDLR 736  
  
RESULT 12  
PCAP\_HUMAN STANDB; 788 AA.  
AC Q96RN5; O15413; Q8NFI6; Q96CT0; Q96IH7; Q9PIT3;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Positive cofactor 2 glutamine/Q-rich-associated protein (PC2  
DE glutamine/Q-rich-associated protein) (TPA-inducible gene-1) (TTG-1)  
DE (Activator-recruited cofactor 105 kDa component) (ARCI05) (CTG repeat  
DE protein 7a).  
GN PCAP OR TIG1 OR ARCI05 OR CTG7A.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCB1\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RC TISSUE=Megakaryocytes, and Placenta;  
RX MEDLINE=20480707; PubMed=11024300;  
RT "A novel glutamine-rich putative transcriptional adaptor protein  
RT (TIG-1), preferentially expressed in placental and bone-marrow  
RT tissues."  
RL Gene 255:389-400(2000).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 2), AND POLYMORPHISM OF POLY-GLN REGION.

RC TISSUE=Cervical carcinoma;  
 RX MEDLINE=21309066; PubMed=11414760;  
 RA Betti L., Mittler G., Przemek G.K.H., Stelzer G., Guenzler B.,  
 RA Amati F., Conti E., Dallapiccola B., Hrabe de Angelis M., Novelli G.,  
 RA Meisterernst M.;  
 RT "Isolation and characterization of a novel gene from the Digeorge  
 RT chromosomal region that encodes for a mediator subunit.";  
 RL Genomics 74:320-332(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC TISSUE=Hepatoma, and Spleen;  
 RA Kawabara A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,  
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;  
 RT "NDO human cDNA sequencing project.";  
 RL Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC TISSUE=Eye, Kidney, and Uterus;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Bosa S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bock S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Pailey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchan J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Buterfield J.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [5]  
 RP SEQUENCE OF 185-573 FROM N.A. (ISOFORM 2).  
 RC TISSUE=Brain cortex;  
 RX MEDLINE=97369492; PubMed=9225980;  
 RA Margolis R.L., Abraham M.R., Gatchell S.B., Li S.-H., Kidwai A.S.,  
 RA Breschel T.S., Strine O.C., Callahan C., McInnis M.G., Ross C.A.;  
 RT "cDNAs with long CAG trinucleotide repeats from human brain.";  
 RL Hum. Genet. 100:114-122(1997).  
 RN [6]  
 RP IDENTIFICATION IN ARC COMPLEX, AND SEQUENCE OF 39-48 AND 525-536.  
 RX MEDLINE=99249346; PubMed=10235267;  
 RA Naeer A.M., Beaurang P.A., Zhou S., Abraham S., Solomon W.B.,  
 RA Tjian R.;  
 RT "Composite co-activator ARC mediates chromatin-directed  
 RT transcriptional activation.";  
 RL Nature 398:828-832(1999).  
 CC -1- FUNCTION: May function as a transcriptional coactivator in RNA  
 CC polymerase II transcription. As a protein complex component it may  
 CC regulate gene transcription, especially following induction of  
 CC protein kinase C activity.  
 CC -1- SUBUNIT: Subunit of the large multiprotein complexes PC2 and  
 CC ARC/DRIP.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q96RNS-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q96RNS-2; Sequence=VSP\_003922;  
 CC -1- TISSUE SPECIFICITY: Expressed in all tissues examined, including  
 CC heart, brain, lung, spleen, thymus, pancreas, blood leukocyte and  
 CC placenta. However, the level of expression varied, with highest  
 CC expression in the placenta and peripheral blood and lowest in the

CC pancreas and kidney.  
 CC -1- INDUCTION: By 12-O-tetradecanoylphorbol-13-acetate (TPA).  
 CC -1- POLYMORPHISM: The poly-Gln region from amino acids 235-262 of  
 CC PCQAP is polymorphic. There are from 15 to 18 repeats in the  
 CC Italian population.  
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to frameshifts  
 CC in positions 13, 600 and 749.  
 CC -1- CAUTION: Ref.3 (BAB85034) sequence differs from that shown due to  
 CC sequencing errors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; AF056191; AAC12944.1; ALT\_FRAME.  
 CC DR EMBL; AF328769; AAK58423.1; -.  
 CC DR EMBL; AK074268; BAB85034.1; ALT\_SEQ.  
 CC DR EMBL; AK090465; BAC03446.1; ALT\_INIT.  
 CC DR EMBL; BC007529; AAH07529.1; -.  
 CC DR EMBL; BC013985; AAH13985.1; -.  
 CC DR EMBL; BC017110; AAH17110.1; ALT\_INIT.  
 CC DR EMBL; U80745; AAB91443.1; -.  
 CC DR Genem: HGNC:14248; PCQAP.  
 CC DR MIM; 607372; -.  
 CC KW Transcription regulation; Activator; Nuclear protein; Polymorphism;  
 CC KW Triplet repeat expansion; Alternative splicing.  
 CC FT DOMAIN 161 174 POLY-GLN.  
 CC FT DOMAIN 178 193 POLY-GLN.  
 CC FT DOMAIN 205 218 POLY-GLN.  
 CC FT DOMAIN 226 239 POLY-GLN.  
 CC FT DOMAIN 243 262 POLY-GLN.  
 CC FT DOMAIN 266 315 PRO-RICH.  
 CC FT DOMAIN 360 367 POLY-GLN.  
 CC FT DOMAIN 449 456 POLY-PRO.  
 CC FT DOMAIN 602 611 POLY-PRO.  
 CC FT VARSPLIC 385 424 Missing (in isoform 2).  
 CC FT VARIANT 261 262 /FtId=VSP\_003922.  
 CC FT CONFLICT 12 12 /FtId=VAR\_01336.  
 CC FT CONFLICT 116 116 S -> R (IN REF. 1).  
 CC FT CONFLICT 154 154 Q -> H (IN REF. 3; BAC03446).  
 CC FT CONFLICT 161 161 Q -> R (IN REF. 3; BAB85034).  
 CC FT CONFLICT 185 186 QO -> EL (IN REF. 5).  
 CC FT CONFLICT 232 287 MISSING (IN REF. 3; BAB85034).  
 CC FT CONFLICT 265 265 O -> E (IN REF. 1 AND 5).  
 CC FT CONFLICT 572 573 IL -> GI (IN REF. 5).  
 CC FT CONFLICT 685 685 L -> V (IN REF. 3).  
 CC SQ SEQUENCE 788 AA; 86753 MW; BB6AC63ED2F97E CRC64;  
 CC  
 CC Query Match 66.1%; Score 37; DB 1; Length 788;  
 CC Best local similarity 59.3%; Pred. No. 57;  
 CC Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC QY 1 KALVQOMELRO 12  
 CC DB 201 QAVVQQQQQLQQ 212  
 CC  
 CC RESULT 13  
 CC SMC2\_CHICK STANDARD; PRT; 1189 AA.  
 CC ID SMC2\_CHICK  
 CC AC Q90988;  
 CC DT 01-NOV-1997 (Rel. 35, Created)  
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 CC DE Structural maintenance of chromosome 2 (Chromosome scaffold protein  
 CC Scf1).

GN SMC2 OR SCII.  
 OS Gallus gallus (Chicken).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 CC Gallus.  
 CC NCBI\_taxid=9031;  
 OK NCBI\_taxid=9031;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fibroblast;  
 RX MEDLINE=95014725; PubMed=7929577;  
 RA Saitoh N., Goldberg I.G., Wood E.R., Earnshaw W.C.;  
 RT "Sci": an abundant chromosome scaffold protein is a member of a  
 RT family of putative Arpases with an unusual predicted tertiary  
 RT structure.";  
 RL J. Cell Biol. 127:303-318(1994).  
 CC -1- FUNCTION: Central component of the condensin complex, a complex  
 CC required for conversion of interphase chromatin into mitotic-like  
 CC condensed chromosomes. The condensin complex probably introduces  
 CC positive supercoils into relaxed DNA in the presence of type I  
 CC topoisomerases and converts nicked DNA into positive knotted forms  
 CC in the presence of type II topoisomerases.  
 CC -1- SUBUNIT: Forms a heterodimer with SMC4. Component of the  
 CC condensin complex, which contains the SMC2 and SMC4 heterodimer,  
 CC and probably some non SMC subunits that regulate the complex.  
 CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic. In interphase  
 CC cells, the majority of the condensin complex is found in the  
 CC cytoplasm, while a minority of the complex is associated with  
 CC chromatin. A subpopulation of the complex however remains  
 CC associated with chromosome foci in interphase cells. During  
 CC mitosis, most of the condensin complex is associated with the  
 CC chromatin. At the onset of prophase, condensin associates with  
 CC chromosome arms and to chromosome condensation. Dissociation from  
 CC chromosomes is observed in late telophase.  
 CC -1- DOMAIN: The flexible hinge domain, which separates the large  
 CC intramolecular coiled coil regions, allows the heterodimerization  
 CC with SMC4, forming a V-shaped heterodimer (by similarity).  
 CC -1- SIMILARITY: Belongs to the SMC family. SMC2 subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL, X80792; CA56767.1; -  
 DR PIR, A54817; A54817.  
 DR InterPro: IPR003405; SMC\_C.  
 DR InterPro: IPR003395; SMC\_N.  
 DR Pfam: PF02483; SMC\_C; 1.  
 DR Pfam: PF02463; SMC\_N; 1.  
 DR Pfam: P000006; ABC\_transporter; 1.  
 KM DNA condensation; Mitosis; Cell cycle; ATP-binding; Coiled coil;  
 KM Nuclear protein.  
 FT NP BIND 32 39 ATP (POTENTIAL).  
 FT DOMAIN 211 503 COILED COIL (POTENTIAL).  
 FT DOMAIN 504 673 FLEXIBLE HINGE.  
 FT DOMAIN 674 1030 COILED COIL (POTENTIAL).  
 FT DOMAIN 1085 1120 ALA/ASP-RICH (DA-BOX).  
 SQ SEQUENCE 1189 AA; 134940 MW; 905EBC89EC45ADSC CRC64;  
 Query Match 64.3%; Score 36; DB 1; Length 1189;  
 Best Local Similarity 58.3%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

RESULT 14  
 HFC1\_MESAU

ID HFC1\_MESAU STANDARD; PRT; 2090 AA.  
 AC P51611;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Host cell factor C1 (HCF) (VP16 accessory protein) (HFC1) (VCAP)  
 DE (CFF).  
 GN HFC1.  
 OS Mesocricetus auratus (Golden hamster).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 CC Mesocricetus.  
 CC NCBI\_taxid=10036;  
 OK NCBI\_taxid=10036;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC Goto H., Nishitani H., Umene K.I., Nakabeppu Y., Nishimoto T.;  
 RA Submitted (FEB-1995) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: UPON LYTIC INFECTION OF PERMISSIVE CELLS, THE HSV  
 CC ACTIVATES VP16 FOR ASSOCIATION WITH THE OCTAMER MOTIF-BINDING  
 CC PROTEIN OCT-1, TO FORM A MULTIPROTEIN-DNA COMPLEX RESPONSIBLE FOR  
 CC ACTIVATING TRANSCRIPTION OF THE HSV IMMEDIATE EARLY GENES (BY  
 CC SIMILARITY).  
 CC -1- SUBUNIT: COMPOSED PREDOMINANTLY OF SIX POLYPEPTIDES RANGING FROM  
 CC 110 TO 150 KDA AND A MINOR 300 KDA POLYPEPTIDE. THE MAJORITY OF N-  
 CC AND C-TERMINAL CLEAVAGE PRODUCTS REMAIN TIGHTLY, ALBEIT  
 CC NONCOVALENTLY, ASSOCIATED (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- DOMAIN: THE HCF REPEAT IS A HIGHLY SPECIFIC PROTEOLYTIC CLEAVAGE  
 CC SIGNAL (BY SIMILARITY).  
 CC -1- PTM: EXISTS AS A FAMILY OF POLYPEPTIDES THAT ARISE THROUGH  
 CC CLEAVAGE AT ONE OR MORE CENTRALLY LOCATED SITES. CLEAVAGE OCCURS  
 CC AT A DEFINED SITE, PCPE/THET, WITHIN THE HCF REPEAT (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: Contains 5 Kelch repeats.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL, D45419; BAA08258.1; -  
 DR InterPro: IPR003961; FN\_III.  
 DR InterPro: IPR006652; Kelch\_rep.  
 DR Pfam, PF00041; FN3; 2.  
 DR Pfam, PF01344; Kelch; 5.  
 DR SMART, SMO0060; FN3; 2.  
 KM Nuclear protein; Kelch repeat; Repeat.  
 FT REPEAT 44 89 KELCH 1.  
 FT REPEAT 93 140 KELCH 2.  
 FT REPEAT 148 194 KELCH 3.  
 FT REPEAT 217 265 KELCH 4.  
 FT REPEAT 266 313 KELCH 5.  
 FT DOMAIN 1010 1448 8 x 26 AA APPROXIMATE REPEATS.  
 FT REPEAT 1010 1035 HCF REPEAT 1.  
 FT REPEAT 1072 1097 HCF REPEAT 2.  
 FT REPEAT 1101 1126 HCF REPEAT 3.  
 FT REPEAT 1157 1182 HCF REPEAT 4. (DEGENERATE).  
 FT REPEAT 1295 1320 HCF REPEAT 5.  
 FT REPEAT 1323 1348 HCF REPEAT 6.  
 FT REPEAT 1358 1383 HCF REPEAT 7. (DEGENERATE).  
 FT REPEAT 1423 1448 HCF REPEAT 8.  
 SQ SEQUENCE 2090 AA; 214942 MW; E495E8B1F2385E17 CRC64;  
 Query Match 64.3%; Score 36; DB 1; Length 2090;  
 Best Local Similarity 63.6%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALVQOMEDLRQ 12  
 ||||| :|||:

Db 1746 ALVQ0000L0E 1756

RESULT 15

YA39\_MYCPN STANDARD; PRT; 338 AA.

AC P75075;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein MPN039 (B01\_orf338).

GN MPN039 OR MP115.

OS Mycoplasma pneumoniae.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI\_TaxId=2104;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 29342 / M129;

RX MEDLINE=97105885; PubMed=8948633;

RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C., Herrmann R.;

RA "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae.";

RT Nucleic Acids Res. 24:4420-4449 (1996).

RL -1 SIMILARITY: BELONGS TO THE MG032 / MG096 / MG288 FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AE000014; AAB95763.1; -

DR PIR; S73441; S73441.

DR InterPro; IPR004319; MG032/096/288\_2.

DR Pfam; PF03086; DUF240; 1.

DR ProDom; PD004834; MG032/096/288\_2; 1.

DR Hypothetical protein; Complete proteome.

SO SEQUENCE 338 AA; 38899 MW; D9D1ABD51B8C7B1 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 338;

Best Local Similarity 50.0%; Pred. No. 56;

Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KALVQOMEQLRQ 12

Db 59 QSLMEQIEQLQQ 70

RESULT 16

YE10\_HAEIN STANDARD; PRT; 394 AA.

AC P44184;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 26-FEB-2003 (Rel. 41, Last annotation update)

DE Protein H11410.

GN H11410.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Haemophilus.

OX NCBI\_TaxId=727;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Rd / KM20 / ATCC 51907;

RX MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kinkness E.F., Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D., Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,

RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhrmann J.L., Geophagen N.S.M., RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., RA Venter J.C.;

RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";

RL Science 269:496-512(1995).

RN [2]

RP IDENTIFICATION BY MASS SPECTROMETRY.

RX MEDLINE=20137488; PubMed=10675023;

RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B., RA Gray C., Fountoulakis M.;

RA "Two-dimensional map of the proteome of Haemophilus influenzae.";

RL Electrophoresis 21:411-429(2000).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U32820; AAC23058.1; -

DR PIR; E64028; E64028.

DR TIGR; H11410; -

DR InterPro; IPR006437; Phage\_term\_2.

DR Pfam; PF04466; G2P; 1.

DR TIGRFAMs; TIGR01547; phage\_term\_2; 1.

DR Complete proteome.

SO SEQUENCE 394 AA; 44782 MW; DA217986BABC772 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 394;

Best Local Similarity 50.0%; Pred. No. 65;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KALVQOMEQLRQ 12

Db 133 KSLMEQIEQLQQ 144

RESULT 17

PURA\_SYNEL STANDARD; PRT; 447 AA.

AC Q8DLG2;

DT 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Adenylosuccinate synthetase (EC 6.3.4.4) (IMP--aspartate ligase) (ADSS) (AMPase).

GN PURA OR T110531.

OS Synecococcus elongatus (Thermosynechococcus elongatus).

OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.

OX NCBI\_TaxId=32046;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BP-1;

RX MEDLINE=22225144; PubMed=12240834;

RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S., RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y., RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N., RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;

RA "Complete genome structure of the thermophilic cyanobacterium Thermosynechococcus elongatus BP-1.";

RT DNA Res. 9:123-130(2002).

RL -1 FUNCTION: Plays an important role in the de novo pathway of purine nucleotide biosynthesis.

CC -1 CATALYTIC ACTIVITY: GTP + IMP + L-aspartate = GDP + phosphate + adenylosuccinate.

CC -1 PATHWAY: AMP biosynthesis; first committed step.

CC -1 SIMILARITY: Belongs to the adenylosuccinate synthetase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP005370; BAC08083.1; -.
CC HAMAP; MF_00011; -.
CC InterPro; IPR001114; Asucc_synthetase.
CC Pfam; PF00709; Adenylysucc_syntc.1.
CC ProDom; PD001188; Asucc_synthetase; 1.
CC TIGRFAMs; TIGR00184; purA; 1.
CC PROSITE; PS01266; ADENYLYSUCCLIN_SYN_1; 1.
CC PROSITE; PS00513; ADENYLYSUCCLIN_SYN_2; 1.
CC Purine biosynthesis; Ligase; GTP-binding; Complete proteome.
CC NP_BIND 12 18
CC ACT_SITE 139 139 BY SIMILARITY.
CC ACT_SITE 146 146 BY SIMILARITY.
CC SEQUENCE 447 AA; 48903 MW; 8F964EB6A89CBE1 CRC64;
SQ
Query Match 62.5%; Score 35; DB 1; Length 447;
Best Local Similarity 50.0%; Pred. No. 73;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLRQ 12
DB 77 KVLLELEQLRQ 88

RESULT 18
ARLY SYN3
ID ARLY SYN3 STANDARD; PRT; 461 AA.
AC P73257.
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Argininosuccinate lyase (EC 4.3.2.1) (Argininosuccinase) (ASAL).
GN ARGH OR SLR1133.
OS Synchocystis sp. (Strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hikosawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- CATALYTIC ACTIVITY: N-(L-arginino)succinate = fumarate + L-
CC arginine.
CC -1- PATHWAY: Arginine biosynthesis; eighth (last) step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the lyase 1 family. Argininosuccinate lyase
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90905; BAA17284.1; -.
CC PIR; S77437; S77437.
CC HSSP; P04424; IAOS.
CC HAMAP; MF_00006; -.

```

```

DR InterPro; IPR003031; D_crystallin.
DR InterPro; IPR000362; Fumarate_lyase.
DR Pfam; PF00206; lyase 1; 1.
DR PRINTS; PR00149; FUMARATELYASE.
DR TIGRFAMs; TIGR00638; argH; 1.
DR PROSITE; PS00163; FUMARATE_LYASES; 1.
CC Arginine biosynthesis; Lyase; Complete proteome.
CC SEQUENCE 461 AA; 50948 MW; B6B8C81B1CFB5FE2 CRC64;
SQ
Query Match 62.5%; Score 35; DB 1; Length 461;
Best Local Similarity 70.0%; Pred. No. 75;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVQQMEQLRQ 12
DB 60 LVQGLEQLRQ 69

RESULT 19
ARLY SYNEL
ID ARLY SYNEL STANDARD; PRT; 463 AA.
AC Q8DLW0.
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Argininosuccinate lyase (EC 4.3.2.1) (Argininosuccinase) (ASAL).
GN ARGH OR TLL0366.
OS Synchococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriyuchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsunoto M., Matsuno A., Nakazaki N.,
RA Shimpō S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
CC -1- CATALYTIC ACTIVITY: N-(L-arginino)succinate = fumarate + L-
CC arginine.
CC -1- PATHWAY: Arginine biosynthesis; eighth (last) step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the lyase 1 family. Argininosuccinate lyase
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP005370; BAC07918.1; -.
CC HAMAP; MF_00006; -.
CC InterPro; IPR003031; D_crystallin.
CC InterPro; IPR000362; Fumarate_lyase.
CC Pfam; PF00206; lyase 1; 1.
CC PRINTS; PR00149; FUMARATELYASE.
CC TIGRFAMs; TIGR00638; argH; 1.
CC PROSITE; PS00163; FUMARATE_LYASES; 1.
CC Arginine biosynthesis; Lyase; Complete proteome.
CC SEQUENCE 463 AA; 51913 MW; A7D26329AAA21412 CRC64;
SQ
Query Match 62.5%; Score 35; DB 1; Length 463;
Best Local Similarity 50.0%; Pred. No. 76;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KALVOQMEQLRQ 12

```





```

FT CONFLICT 381 381 L -> I (IN REF. 1)
FT CONFLICT 407 407 I -> Y (IN REF. 1)
FT CONFLICT 670 670 V -> A (IN REF. 1)
FT CONFLICT 670 670 V -> G (IN REF. 1)
FT CONFLICT 710 710 I -> V (IN REF. 3)
SQ SEQUENCE 733 AA; 80908 MW; C15B4C9EBB3333C5 CRC64;

Query Match
Best Local Similarity 72.7%; Score 35; DB 1; Length 733;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KALVOOMEQLR 11
Db 696 KALSNOMEQAR 706

RESULT 22
KFSC_MOUSE STANDARD; PRT; 956 AA.
ID KFSC_MOUSE
AC P28738; Q922F8;
DT 01-DEC-1992 (Rel. 24, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Kinesin heavy chain isoform 5C (Kinesin heavy chain neuron-specific
2).
GN KIF5C OR NKHC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP PRELIMINARY SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Brain;
RA Kato K.;
RT "A collection of cDNA clones with specific expression patterns in
mouse brain."
RL Eur. J. Neurosci. 2:704-711(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9900842; PubMed=9782088;
RA Xia C., Rahman A., Yang Z., Goldstein L.S.B.;
RT "Chromosomal localization reveals three kinesin heavy chain genes in
mouse."
RL Genomics 52:209-213(1998).
CC -1- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT.
CC -1- SUBUNIT: OLIGOMER COMPOSED OF TWO HEAVY CHAINS AND TWO LIGHT
CC CHAINS.
CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A LARGE GLOBULAR N-
CC TERMINAL DOMAIN WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY OF
CC KINESIN (IT HYDROLYZES ATP AND BINDS MICROTUBULE), A CENTRAL
CC ALPHA-HELICAL COILED COIL DOMAIN THAT MEDIATES THE HEAVY CHAIN
CC DIMERIZATION, AND A SMALL GLOBULAR C-TERMINAL DOMAIN WHICH
CC INTERACTS WITH OTHER PROTEINS (SUCH AS THE KINESIN LIGHT CHAINS),
CC VESICLES AND MEMBRANOUS ORGANELLES.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC SUBFAMILY.
CC -1- SUBUNIT: OLIGOMER COMPOSED OF TWO HEAVY CHAINS AND TWO LIGHT
CC CHAINS.
CC -1- CAUTION: REF.1 SEQUENCE SEEMS TO HAVE BEEN A HYBRID OF A KIF5A
CC AND A KIF5C SEQUENCE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X61435; CAA43677.1; -
CC EMBL, AF067180; AAC79804.1; -
CC DR HSSP; P56536; 2KIN.
CC MGD; MGI:1098269; Kif5c.

```

```

DR GO; GO:0005515; F:protein binding activity; IPI.
DR GO; GO:0008045; P:motor axon guidance; IMP.
DR InterPro; IPR001752; Kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
FT DOMAIN 1 386 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 406 923 COILED COIL.
FT DOMAIN 859 956 GLOBULAR.
FT DOMAIN 174 315 MICROTUBULE-BINDING.
FT NP BIND 86 93 ATP (BY SIMILARITY).
SQ SEQUENCE 956 AA; 109240 MW; D5A6C701A2911AB2 CRC64;

Query Match
Best Local Similarity 66.7%; Score 35; DB 1; Length 956;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KALVOOMEQLRQ 12
Db 700 KALSNOMESHRE 711

RESULT 23
KFSC_HUMAN STANDARD; PRT; 957 AA.
ID KFSC_HUMAN
AC O60282; Q95079;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Kinesin heavy chain isoform 5C (Kinesin heavy chain neuron-specific
2).
GN KIF5C OR NKHC2 OR KIAA0531.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98290545; PubMed=9628581;
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro."
RL DNA Res. 5:31-39(1998).
RN [2]
RP SEQUENCE OF 355-585 FROM N.A.
RA Engländer S., Sharp A.H., Colomer V., Tokito M.K., Lanahan A.,
RA Worley P., Holzbaur E.L.F., Ross C.A.;
RT "Huntingtin associated protein 1 (HAP1) interacts with the p150glcd
RT subunit of dynein."
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT.
CC -1- SUBUNIT: OLIGOMER COMPOSED OF TWO HEAVY CHAINS AND TWO LIGHT
CC CHAINS.
CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN BRAIN, PROSTATE, AND
CC TESTIS, AND MODERATE EXPRESSION IN KIDNEY, SMALL INTESTINE, AND
CC OVARY.
CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A LARGE GLOBULAR N-
CC TERMINAL DOMAIN WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY OF
CC KINESIN (IT HYDROLYZES ATP AND BINDS MICROTUBULE), A CENTRAL
CC ALPHA-HELICAL COILED COIL DOMAIN THAT MEDIATES THE HEAVY CHAIN
CC DIMERIZATION, AND A SMALL GLOBULAR C-TERMINAL DOMAIN WHICH
CC INTERACTS WITH OTHER PROTEINS (SUCH AS THE KINESIN LIGHT CHAINS),
CC VESICLES AND MEMBRANOUS ORGANELLES.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC SUBFAMILY.
CC -----

```



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; AB011103; BAA25457.1; -  
 CC EMBL; AF010146; AAD01436.1; -  
 CC HSSP; P56536; 2KIN.  
 CC GeneW; HGNC:6325; KIF5C.  
 CC MIM; 604593; -  
 CC GO; GO:0005871; C:kinesin complex; TAS.  
 CC GO; GO:0003777; F:microtubule motor activity; TAS.  
 CC GO; GO:0006966; P:organelle organization and biogenesis; TAS.  
 CC InterPro; IPR001752; kinesin\_motor.  
 CC Pfam; PF00225; Kinesin; 1.  
 CC PRINTS; PR00380; KINESINHEAVY.  
 CC SMART; SM00129; KISC; 1.  
 CC PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
 CC PROSITE; PS00067; KINESIN\_MOTOR\_DOMAIN2; 1.  
 CC Motor protein; Microtubules; ATP-binding; Coiled coil.  
 CC FT DOMAIN 1 386 KINESIN\_MOTOR (BY SIMILARITY).  
 CC FT DOMAIN 406 923 COILED COIL.  
 CC FT DOMAIN 859 956 MICROTUBULE-BINDING.  
 CC FT NP\_BIND 174 315 ATP (BY SIMILARITY).  
 CC FT CONFLICT 355 360 TIKANI -> STIASV (IN REF. 2).  
 CC FT CONFLICT 583 585 EFT -> DRV (IN REF. 2).  
 CC SQ SEQUENCE 957 AA; 109494 MW; AAF25BBLIC94322A CRC64;

Query Match 62.5%; Score 35; DB 1; Length 957;  
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KALVQOMEQLRQ 12  
 DB 701 KALVQOMESHRE 712

RESULT 24  
 ID MYHD\_HUMAN STANDARD; PRT; 1938 AA.  
 AC Q9UKX3; Q95252;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin heavy chain, skeletal muscle, extraocular (MYHC-eo).  
 GN MYH3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID:9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Extraocular muscle;  
 RX MEDLINE=99318869; PubMed=10388558;  
 RA Weiss A., Schiaffino S., Leinwand L.A.;  
 RT "Comparative sequence analysis of the complete human sarcomeric myosin  
 RT heavy chain family: implications for functional diversity.";  
 RL J. Mol. Biol. 290:61-75(1999).  
 [2]  
 RP SEQUENCE OF 1917-1938 FROM N.A.  
 RC TISSUE=Extraocular muscle;  
 RX MEDLINE=99026150; PubMed=9806854;  
 RA Winters L.M., Briggs M.M., Schachat F.;  
 RT "The human extraocular muscle myosin heavy chain gene (MYH13) maps to  
 RT the cluster of fast and developmental myosin genes on chromosome 17.";  
 RL Genomics 54:188-189(1998).  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)

CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.  
 CC -1- SIMILARITY: Contains 1 IQ domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; AF111782; AAD29948.1; -  
 CC EMBL; AF075248; AAC83241.1; -  
 CC HSSP; P13538; 2MYS.  
 CC GeneW; HGNC:7571; MYH13.  
 CC MIM; 603487; -  
 CC GO; GO:0005859; C:muscle myosin; TAS.  
 CC GO; GO:0003779; F:actin binding activity; NAS.  
 CC GO; GO:0005524; F:ATP binding activity; NAS.  
 CC GO; GO:0005516; F:calmodulin binding activity; NAS.  
 CC GO; GO:0003776; F:muscle motor activity; TAS.  
 CC InterPro; IPR000048; IQ\_region.  
 CC InterPro; IPR001609; myosin\_head.  
 CC InterPro; IPR004009; Myosin\_N.  
 CC InterPro; IPR002928; Myosin\_tail.  
 CC Pfam; PF00612; IQ; 2.  
 CC Pfam; PF00063; myosin\_head; 1.  
 CC Pfam; PF02736; myosin\_N; 1.  
 CC Pfam; PF01576; myosin\_tail; 1.  
 CC PRINTS; PR00193; MYOSINHEAVY.  
 CC PRODOM; PD000355; myosin\_head; 1.  
 CC SMART; SM00015; IQ; 1.  
 CC SMART; SM00242; MYSC; 1.  
 CC PROSITE; PS50096; IQ; 1.  
 CC MYosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 CC Calmodulin-binding; ATP-binding; Methylation; Alkylation;  
 CC MultiGene family.  
 CC FT DOMAIN 1 784 MYOSIN HEAD-LIKE.  
 CC FT DOMAIN 785 814 IQ.  
 CC FT NP\_BIND 179 186 ATP (POTENTIAL).  
 CC FT DOMAIN 659 681 ACTIN-BINDING (BY SIMILARITY).  
 CC FT DOMAIN 761 775 ACTIN-BINDING (BY SIMILARITY).  
 CC FT MOD\_RES 130 130 METHYLATION (TRI-) (POTENTIAL).  
 CC FT MOD\_RES 699 699 ALKYLATION (SH-1) (POTENTIAL).  
 CC FT MOD\_RES 709 709 ALKYLATION (SH-2) (POTENTIAL).  
 CC SQ SEQUENCE 1938 AA; 223678 MW; 1F6D006416381CD5 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 1938;  
 Best Local Similarity 50.0%; Pred. No. 3.1e+02;  
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 KALVQOMEQLRQ 12  
 DB 1310 QALTQOLEELKR 1321

RESULT 25  
 ID BIR6\_HUMAN STANDARD; PRT; 4829 AA.  
 AC Q9NR09; Q9ULD1;  
 DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Baculoviral IAP repeat-containing protein 6 (ubiquitin-conjugating  
 DE BIR-domain enzyme apollon).  
 CN BIRC6 OR KIAA1289.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20012759; PubMed=10544019;  
 RA Chen Z., Naito M., Hori S., Mashima T., Tsuruo T.;  
 RA "A human IAP-family gene, apollon, expressed in human brain cancer  
 RA cells";  
 RL Biochem. Biophys. Res. Commun. 264:847-854(1999).  
 RN [2]  
 RP SEQUENCE OF 3238-4829 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20039619; PubMed=10574462;  
 RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,  
 RA Ohara O.;  
 RA "Prediction of the coding sequences of unidentified human genes. XV.  
 RA The complete sequences of 100 new cDNA clones from brain which code  
 RA for large proteins in vitro.";  
 RL DNA Res. 6:337-345(1999).  
 CC -1- FUNCTION: MAY PROTECT CELLS FROM UNDERGOING APOPTOSIS.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN CANCER CELLS.  
 CC -1- SIMILARITY: Contains 1 BIR repeat.  
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE UBIQUITIN-  
 CC CONJUGATING ENZYME FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL, AF265555; AAF75772.1; -;  
 CC EMBL, AB033115; BAA86603.1; -;  
 CC HSSP; Q13490; IOBH.  
 CC Genew; HGNC:13516; BIRC6.  
 CC MIM; 605638; -;  
 CC GO; GO:0004840; F-ubiquitin conjugating enzyme activity; TAS.  
 CC GO; GO:0008916; Protein-apoptosis; TAS.  
 CC InterPro; IPR001370; BIR.  
 CC InterPro; IPR000608; UBQ\_conjugat.  
 CC Pfam; PF00653; BIR; 1.  
 CC Pfam; PF00179; UQ\_con; 1.  
 CC ProDom; PD000461; UBQ\_conjugat; 1.  
 CC SMART; SM00238; BIR; 1.  
 CC SMART; SM00212; UBQC; 1.  
 CC PROSITE; PS01282; BIR\_REPEAT\_1; FALSE\_NEG.  
 CC PROSITE; PS00143; BIR\_REPEAT\_2; 1.  
 CC PROSITE; PS00183; UBIQUITIN\_CONJUGAT\_1; FALSE\_NEG.  
 CC PROSITE; PS50127; UBIQUITIN\_CONJUGAT\_2; 1.  
 CC Apoptosis; Thiol protease inhibitor; Ubl conjugation pathway; Ligase.  
 KM APOPTOSIS; Thiol protease inhibitor; Ubl conjugation pathway; Ligase.  
 FT REPEAT 256 330 BIR.  
 FT DOMAIN 4548 4676 UBIQUITIN-CONJUGATING.  
 FT BINDING 4597 4597 UBIQUITIN (BY SIMILARITY).  
 FT DOMAIN 2 8 POLY-ALA.  
 FT DOMAIN 1632 1640 POLY-ALA.  
 SQ SEQUENCE 4829 AA; 527604 MW; C67126A672CD3653 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 4829;  
 Best Local Similarity 77.8%; Pred. No. 7.5e+02;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVQOMEQ 9  
 |||||:|||||;

Db 2222 KALVQOMEK 2230

RESULT 26  
 YBGL\_HAEIN STANDARD; PRT; 213 AA.  
 AC P44299;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein H1731.  
 GN H1731.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.  
 OX NCBI\_TaxId=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Rd / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-T., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,  
 RA Gnehm C.V., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 RA Rd.";  
 RL Science 269:496-512(1995).  
 CC -1- SIMILARITY: STRONG, TO E. COLI YBGL.  
 CC -1- SIMILARITY: TO B. SUBTILIS YCSJ AND YEAST UREA AMIDOLYASE (DUR1.2).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL, U32845; AAC23377.1; -;  
 CC PIR; C64041; C64041.  
 CC TIGR; H11731; -;  
 CC InterPro; IPR003833; DUF213.  
 CC Pfam; PF02682; AHS1; 1.  
 CC Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 213 AA; 23641 MW; 1A2360FC83F2D77E CRC64;

Query Match 60.7%; Score 34; DB 1; Length 213;  
 Best Local Similarity 70.0%; Pred. No. 53;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KALVQOMEQ 10  
 |||||:|||||;

Db 61 KPLVORLQL 70

RESULT 27  
 CBIK\_SALTY STANDARD; PRT; 264 AA.  
 ID CBIK\_SALTY  
 AC 005552;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cobalt chelatase (EC 4.99.1.-).  
 GN CBIK OR STM2025.  
 OS Salmonella typhimurium.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Salmonella.  
 OX NCBI\_TaxId=602;

```

RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF N-TERMINUS.
RC STRAIN=LT2;
RX MEDLINE=93273696; PubMed=8501034;
RA Roth J.R., Lawrence J.G., Rubenfield M., Kieffer-Higgins S.,
RA Church G.M.;
RT "Characterization of the cobalamin (vitamin B12) biosynthetic genes
RL of Salmonella typhimurium."
RN J. Bacteriol. 175:3303-3316(1993).
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Malyan E.,
RA Ryan E., Sun H., Florea L., Miller W., Stenking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RL LT2."
RN Nature 413:852-856(2001).
RP CHARACTERIZATION.
RX MEDLINE=97294459; PubMed=9150215;
RA Raux E., Thernes C., Heathcote P., Rambach A., Warren M.J.;
RT "A role for Salmonella typhimurium cbkX in cobalamin (vitamin B12) and
RL siroheme biosynthesis."
RN J. Bacteriol. 179:3202-3212(1997).
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=99379809; PubMed=10451360;
RA Schubert H.L., Raux E., Wilson K.S., Warren M.J.;
RT "Common chelate design in the branched tetrahydropyrole pathways of
RL heme and anaerobic cobalamin synthesis."
RN Biochemistry 38:10660-10669(1999).
CC -I- PATHWAY: Cobalamin biosynthesis.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, L12006; AAA27262.1; -.
DR EMBL, AE008789; AAL20929.1; -.
DR PDB, 1QGO; 23-SEP-99.
DR StyGene; SG10043; cbkX.
KW Cobalamin biosynthesis; Lyase; 3D-structure; Complete proteome.
CC
FT STRAND 3
FT HELIX 15
FT TURN 32
FT TURN 34
FT STRAND 37
FT STRAND 42
FT HELIX 45
FT HELIX 55
FT TURN 62
FT TURN 72
FT STRAND 76
FT STRAND 81
FT HELIX 85
FT HELIX 88
FT HELIX 101
FT STRAND 106
FT STRAND 109
FT STRAND 114
FT HELIX 117
FT TURN 129
FT TURN 136
FT STRAND 137
FT STRAND 138
FT STRAND 144
FT HELIX 149
FT TURN 166
FT STRAND 169
FT TURN 173
FT TURN 175

```

```

FT HELIX 180 190
FT TURN 191 191
FT STRAND 194 199
FT HELIX 206 209
FT TURN 210 213
FT TURN 217 218
FT HELIX 220 226
FT TURN 227 228
FT STRAND 231 233
FT HELIX 238 240
FT HELIX 242 257
SQ SEQUENCE 264 AA; 29238 MW; 34667EDACF603BEC CRC64;

Query Match 60.7%; Score 34; DB 1; Length 264;
Best Local Similarity 66.7%; Pred. No. 65;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KALVQMEQLRQ 12
Db 125 QALRQMPSLRQ 136

```

```

RESULT 28
SECA_BACFI STANDARD; PRT; 473 AA.
AC P96313;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last Sequence Update)
DT 30-MAY-2000 (Rel. 39, Last Annotation Update)
DE Preprotein translocase seca subunit (Fragment).
GN SECA.
OS Bacillus firmus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1159;
RN [1]
RP SEQUENCE FROM N.A.
RA Bauer M., Baerentzen E.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: INVOLVED IN PROTEIN EXPORT. INTERACTS WITH THE SECY/SECE
CC SUBUNITS. SECA HAS A CENTRAL ROLE IN COUPLING THE HYDROLAYSIS OF
CC ATP TO THE TRANSFER OF PRE-SECRETORY PERIPLASMIC AND OUTER
CC MEMBRANE PROTEINS ACROSS THE MEMBRANE (BY SIMILARITY).
CC -I- SUBUNIT: PART OF THE PROKARYOTIC PROTEIN TRANSLLOCATION APPARATUS
CC WHICH COMPRISE SECA, SECB, SECD, SECE, SECF, SECG AND SECY
CC (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF PLASMA MEMBRANE
CC (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE SECA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X9401; CA6777.1; -.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004027; SEC_C_motif.
DR InterPro; IPR000185; SECA.
DR Pfam; PF00271; helicase_C; 1.
DR Pfam; PF02810; SEC-C; 1.
DR Pfam; PF01043; SecA_protein; 1.
DR PROSITE; PS01312; SECA; 1.
KW Protein transport; ATP-binding; Membrane; Translocation; Transport.
FT NON TER 1
SQ SEQUENCE 473 AA; 54466 MW; F229F9D0E1B4234 CRC64;

Query Match 60.7%; Score 34; DB 1; Length 473;
Best Local Similarity 66.7%; Pred. No. 1,2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      4 VOQMEQLRQ 12
      : ||:||||
DB      363 IDQMDQLRQ 371

RESULT 29
SESI_MOUSE STANDARD; PRT; 492 AA.
ID SESI_MOUSE
AC P58006;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
SEstrin 1 (p53-regulated protein PA26).
SESTR1 OR PA26.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP RECONSTRUCTION FROM ESTS.
RX MEDLINE=22578171; PubMed=12607115;
RA Peeters H., Debeer P., Bairoch A., Milquet V., Huysmans C.,
RA Parhoens E., Frys J.P., Gewillig M., Nakamura Y., Nikawa N.,
RA Van De Ven W., Derynck K.,
RT "PA26 is a candidate gene for heterotaxia in humans: identification of
RT a novel PA26-related gene family in human and mouse."
RL Hum. Genet. 112:573-580(2003).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the seestrin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AM121582; -; NOT ANNOTATED CDS.
CC EMBL, BE913206; -; NOT ANNOTATED CDS.
CC EMBL, BG088667; -; NOT ANNOTATED CDS.
CC MGD; MGI:2155278; Pa26.
CC InterPro; IPR006730; PA26.
CC Pfam; PF04636; PA26; 1.
CC Nuclear protein.
CC FT CONFLICT 301 301 F -> L (IN BG088667).
CC SEQUENCE 492 AA; 56632 MW; BE4156C5DAD7A8D8 CRC64;

Query Match 60.7%; Score 34; DB 1; Length 492;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      1 KALVOQMEQLRQ 12
      : ||:||||
DB      274 EALVQMEQRLQ 285

RESULT 30
SADI_SCHPO STANDARD; PRT; 514 AA.
ID SADI_SCHPO
AC Q09925;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Spindle pole body associated protein sadi.
GN SADI OR SPBC12D12.01 OR SPBC145.01C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxId=4896;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=95263659; PubMed=7744953;
RA Hagan I., Yanagida M.;
RT "The product of the spindle formation gene sadi+ associates with the
RT fission yeast spindle pole body and is essential for viability."
RL J. Cell Biol. 129:1033-1047(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
RA Skelton J., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymoprez B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaute V., Motier S.,
RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -1- FUNCTION: ASSOCIATES WITH THE SPINDLE POLE BODY AND MAY MAINTAIN
CC A FUNCTIONAL INTERFACE BETWEEN THE NUCLEAR MEMBRANE AND THE
CC MICROTUBULE MOTOR PROTEINS OR MAY PROVIDE AN ANCHOR FOR THESE
CC MOTOR PROTEINS. ESSENTIAL FOR VIABILITY.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. NUCLEAR ENVELOPE.
CC -1- SIMILARITY: Contains 1 Unc84 (SUN) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X85105; CAAS9426.1; -.
CC EMBL, AL035085; CAA22675.1; -.
CC EMBL, AL022104; CAA17899.1; -.
CC PIR; A57280; A57280.
CC Genedb_Spombe; SPBC12D12.01; -.
CC Cell division; Microtubules; Mitosis; Transmembrane; Nuclear protein;
CC Phosphorylation.
CC KW TRANSMEM 170 188 POTENTIAL.
CC FT DOMAIN 373 489 UNCR4.
CC FT MOD_RES 236 236 PHOSPHORYLATION (BY CDC2) (POTENTIAL).
CC SEQUENCE 514 AA; 58077 MW; DCA4CAC1EF98F0D CRC64;

Query Match 60.7%; Score 34; DB 1; Length 514;
Best Local Similarity 58.3%; Pred. No. 1.3e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 KALVOQMEQLRQ 12
      : ||:||||
DB      261 KILQKVEQLRQ 272

RESULT 31
PRPR_SALTY

```

ID PRPR\_SALTY STANDARD; PRT; 541 AA.  
AC P74839;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Prodigate catabolism operon regulatory protein.  
PRPR OR STM0367.  
GN Salmonella typhimurium.  
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI\_TaxID=602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LT2;  
RX MEDLINE=97158691; PubMed=9006051;  
RA Horswill A.R., Escalante-Semerena J.C.;  
RT "Prophage catabolism in Salmonella typhimurium LT2: two divergently  
transcribed units comprise the prp locus at 8.5 centisomes, prpR  
encodes a member of the sigma-54 family of activators, and the  
prpBCDE genes constitute an operon."  
RT prpBCDE genes constitute an operon."  
RL J. Bacteriol. 179:928-940(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;  
RX MEDLINE=21534948; PubMed=11677609;  
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
Ryan B., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
Waterston R., Wilson R.K.;  
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
LT2."  
RL Nature 413:852-856(2001).  
CC -1- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF THE  
PROPHAGE CATABOLISM OPERON.  
CC -1- SIMILARITY: Contains 1 sigma-54 factor interaction ATP-binding  
domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U51879; AAC4813.1; -  
CC EMBL; AE008712; AAL19321.1; -  
CC STyGene; SG10690; PRPR.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR002197; HTH\_Fis.  
DR InterPro; IPR002078; S1954\_Interact.  
DR Pfam; PF02954; HTH\_8; 1.  
DR Pfam; PF00158; Sigma54\_activat; 1.  
DR PRINTS; PRO1590; HTHFIS.  
DR SMART; SMO0382; AAA; 1.  
DR TIGRfams; TIGR01199; HTH\_fis; 1.  
DR PROSITE; PS00675; SIGMA54\_INTERACT\_1; FALSE\_NEG.  
DR PROSITE; PS00676; SIGMA54\_INTERACT\_2; 1.  
DR PROSITE; PS00688; SIGMA54\_INTERACT\_3; 1.  
DR PROSITE; PS00445; SIGMA54\_INTERACT\_4; 1.  
KW Transcription regulation; DNA-binding; ATP-binding; Complete proteome.  
FT DOMAIN 221 464 SIGMA-54 FACTOR INTERACTION (POTENTIAL).  
FT NP\_BIND 321 330 ATP (POTENTIAL).  
FT DN\_BIND 513 532 H-T-H MOTIF (BY SIMILARITY).  
SQ SEQUENCE 541 AA; 60338 MW; 6E231219FCFEDC6 CRC64;

Query Match 60.7%; Score 34; DB 1; Length 541;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

6 QMEQOLRQ 12  
|||||||

DB 227 QMEQOLRQ 233  
RESULT 32  
APB3 RAT STANDARD; PRT; 569 AA.  
ID APB3 RAT  
AC O70248;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Amyloid beta A4 precursor protein-binding family A member 3 (Neuron-  
specific X11L2 protein) (Neuronal Munc18-1-interacting protein 3)  
DE (Mint-3) (Adaptor protein X11gamma).  
GN APB3 OR MINT3.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99075474; PubMed=9860131;  
RA Okamoto M., Suedhof T.C.;  
RT "Mint 3: a ubiquitous mint isoform that does not bind to munc18-1 or  
-2."  
RL Eur. J. Cell Biol. 77:161-165(1998).  
CC -1- FUNCTION: May modulate processing of the beta-amyloid precursor  
protein (APP) and hence formation of beta-APP.  
CC -1- SUBUNIT: Binds to the cytoplasmic domain of amyloid protein (APP).  
CC -1- TISSUE SPECIFICITY: Ubiquitous.  
CC -1- DOMAIN: Composed of an N-terminal domain, a middle  
phosphotyrosine-binding domain (PID/PTB) that mediates binding  
with the cytoplasmic domain of the beta-amyloid precursor protein,  
and two C-terminal PDZ domains thought to attach proteins to the  
plasma membrane.  
CC -1- SIMILARITY: Contains 2 PDZ/DHR domains.  
CC -1- SIMILARITY: Contains 1 PID domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF029109; AAC17978.1; -  
CC HSSP; Q02410; IAQC.  
DR InterPro; IPR001478; PDZ.  
DR InterPro; IPR006020; PTB\_PID.  
DR Pfam; PF00595; PDZ; 2.  
DR Pfam; PF00640; PID; 1.  
DR SMART; SMO0228; PDZ; 2.  
DR SMART; SMO0462; PTB; 1.  
DR PROSITE; PS0106; PDZ; 1.  
DR PROSITE; PS0106; PDZ; 2.  
DR PROSITE; PS01179; PID; 1.  
KW Protein transport; Repeat.  
FT DOMAIN 212 376 PID.  
FT DOMAIN 389 475 PDZ 1.  
FT DOMAIN 480 554 PDZ 2.  
SQ SEQUENCE 569 AA; 60639 MW; 27F9D55A4D2B939C CRC64;

Query Match 60.7%; Score 34; DB 1; Length 569;  
Best Local Similarity 70.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 KALVQMEQL 10  
:|||||:  
58 KALVQMEQL 67

RESULT 33  
CAO2\_CANTR STANDARD; PRT; 708 AA.  
ID CAO2\_CANTR

AC P06598; 01-JAN-1988 (Rel. 06, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Acyl-coenzyme A oxidase II (EC 1.3.3.6) (Acyl-CoA oxidase) (EXP-4)  
 DE (Peroxisomal fatty acyl-CoA oxidase).  
 GN AOX OR FOX-4.  
 OS Candida tropicalis (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxId=5482;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 20336 / PK233;  
 RX MEDLINE=87248070; PubMed=3596241;  
 RA Murray W.W., Rachubinski R.A.;  
 RT "The primary structure of a peroxisomal fatty acyl-CoA oxidase from  
 RT the yeast Candida tropicalis pk233.";  
 RL Gene 51:119-128 (1987).  
 RN [2]  
 RP REVISIONS TO 366; 380 AND 384.  
 RA Murray W.W., Rachubinski R.A.;  
 RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 20336 / PK233;  
 RX MEDLINE=86149279; PubMed=3456583;  
 RA Okazaki K., Takeuchi T., Kambara N., Fukui S., Kubota I., Kamiryo T.;  
 RT "Two acyl-coenzyme A oxidases in peroxisomes of the yeast Candida  
 RT tropicalis: primary structures deduced from genomic DNA sequence.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:1232-1236 (1986).  
 RN [4]  
 RP SEQUENCE OF 208-709 FROM N.A.  
 RC STRAIN=RRI;  
 RX MEDLINE=87280361; PubMed=3611187;  
 RA Small G.M., Lazatow P.B.;  
 RT "Import of the carboxy-terminal portion of acyl-CoA oxidase into  
 RT peroxisomes of Candida tropicalis.";  
 RL J. Cell Biol. 105:247-250 (1987).  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + O(2) = trans-2,3-dehydroacyl-CoA +  
 CC H(2)O(2).  
 CC -1- COFACTOR: FAD.  
 CC -1- PATHWAY: Peroxisomal fatty acid beta-oxidation system; first step.  
 CC -1- SUBUNIT: Homooctamer.  
 CC -1- SUBCELLULAR LOCATION: Peroxisomal.  
 CC -1- SIMILARITY: BELONGS TO THE ACYL-COA OXIDASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M16193; AAA34332.2; -;  
 CC DR EMBL; M12160; AAA34362.1; -;  
 CC DR EMBL; Y00623; CAAG68660.1; -;  
 CC DR EMBL; Y00623; CAAG68661.1; ALT\_INIT.  
 CC DR EMBL; Y00623; CAAG68662.1; ALT\_INIT.  
 CC DR PIR; A29047; OXCKX.  
 CC DR InterPro; IPR002655; ACOX.  
 CC DR InterPro; IPR006090; Acyl-CoA\_dh\_C.  
 CC DR Pfam; PF01756; ACOX; 1.  
 CC DR Pfam; PF00441; Acyl-CoA\_dh; 1.  
 CC DR Oxidoreductase; Fatty acid metabolism; Peroxisome; Flavoprotein; FAD;  
 CC MultiGene family.  
 KW INIT MET  
 FT 0  
 FT CONFLICT 216 216 O -> E (IN REF. 4).  
 FT CONFLICT 245 245 P -> A (IN REF. 3 AND 4).  
 FT CONFLICT 335 335 N -> K (IN REF. 3).  
 FT CONFLICT 358 353 FVPMSPSPVPSRLNTPSPSPMSNMTSPKRTTPRL -> L  
 AAAYISAGALKEVDIHTHTLAELDAVAEKNDTKA (IN

FT FT REF. 3 AND 4).  
 FT CONFLICT 435 435 H -> Y (IN REF. 3).  
 FT CONFLICT 462 462 G -> A (IN REF. 3).  
 FT CONFLICT 495 495 E -> S (IN REF. 3).  
 FT CONFLICT 576 578 ELA -> DLV (IN REF. 3).  
 FT CONFLICT 697 697 Q -> E (IN REF. 3 AND 4).  
 SQ SEQUENCE 708 AA; 78962 MW; FE279F7FEB9A9E5B CRC64;  
 Query Match 60.7%; Score 34; DB 1; Length 708;  
 Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 KALVQOMEQ 9  
 Db 48 KALVQOMER 56  
 ID DECI\_YEAST STANDARD; PRT; 796 AA.  
 AC 012387;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE DECI protein (MDM20 protein).  
 GN DECI OR MDM20 OR YOL076W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxId=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kingsbury T.J., Hoyt M.A.;  
 RL Submitted (SEP-1995) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Hermann G.J., Shaw J.M.;  
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97321807; PubMed=9178509;  
 RA Tzeremia M., Katsoulou C., Alexandraki D.;  
 RT "Sequence analysis of a 33.2 Kb segment from the left arm of yeast  
 RT chromosome XV reveals eight known genes and ten new open reading  
 RT frames including homologues of ABC transporters, inositol  
 RT phosphatases and human expressed sequence tags.";  
 RL Yeast 13:583-589 (1997).  
 CC -1- FUNCTION: GENETICALLY INTERACTS WITH CIN8.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U36382; AAA9781.1; -;  
 CC DR EMBL; U54799; AAB00196.1; -;  
 CC DR EMBL; Z74818; CAA99086.1; -;  
 CC DR PIR; S66769; S66769.  
 CC DR SGD; S0005436; MDM20.  
 CC DR GO; GO:0005622; C:intracellular; IDA.  
 CC DR GO; GO:0008580; F:cytoskeleton regulator activity; IGT.  
 CC DR GO; GO:0007010; P:cytoskeleton organization and biogenesis; IMP.  
 CC DR GO; GO:0000001; P:mitochondrion inheritance; IGI.  
 SQ SEQUENCE 796 AA; 92808 MW; 30E57DB1B12DCD2C CRC64;  
 Query Match 60.7%; Score 34; DB 1; Length 796;  
 Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 KALVQOMEQLRQ 12

```

Db      751 KALVOAMEBLK 762

RESULT 35
GCR3_YEAST STANDARD; PRT; 861 AA.
ID GCR3_YEAST
AC P34160;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE GCR3 protein (STOI protein) (STOI protein).
GN GCR3 OR STOI OR YMR125W OR YMR564.07 OR YMR553.01.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;

SEQUENCE FROM N.A.
RP MEDLINE=92380925; PubMed=1512188;
RA Uemura H., Jigami Y.;
RT "GCR3 encodes an acidic protein that is required for expression of
RL glycolytic genes in Saccharomyces cerevisiae.";
RN J. Bacteriol. 174:5526-5532(1992).
[2]
SEQUENCE FROM N.A.
RP Pandit S., Sternlanz R.;
RA Submitted (DEC-1992) to the EMBL/Genbank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RP Guo Z., Russo P., Sherman F.;
RA Submitted (XXX-1994) to the EMBL/Genbank/DBJ databases.
RN [4]
SEQUENCE FROM N.A.
RP STRAIN=S288c / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RL XII."
RN Nature 387:90-93(1997).
CC -1- FUNCTION: REQUIRED FOR EXPRESSION OF GLYCOLYTIC GENES. HAS
CC CERTAIN CHARACTERISTICS OF A TRANSCRIPTIONAL ACTIVATOR.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: SOME, TO HUMAN CBP80.
CC -1- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 708
CC ONWARD AND IS SHORTER (725 AA) DUE TO A FRAMESHIFT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10224; BA01076.1; ALT SEQ.
DR EMBL; L07650; -; NOT ANNOTATED CDS.
DR EMBL; L27744; -; NOT ANNOTATED CDS.
DR EMBL; Z49273; CA889274.1; -
DR EMBL; Z48622; CA88550.1; -
DR PIR; A44919; A44919.
DR SGD; S0004732; STOI.
DR GO; GO:0000243; C:commitment complex; IPI.
DR GO; GO:0005846; C:snRNA cap binding complex; IDA.
DR GO; GO:0003729; F:mRNA binding activity; IPI.
DR GO; GO:0006371; F:mRNA binding activity; IPI.
DR InterPro; IPR003890; IF_eIF4G.
DR Pfam; PF02854; MIF4G; 1.
DR SMART; SM00543; MIF4G; 1.
DR DNA-binding; Nuclear protein.
KW

```

```

FT DOMAIN 22 30 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 774 801 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 802 825 ARG/LYS-RICH (BASIC).
FT CONFLICT 164 164 D -> V (IN REF. 3).
FT CONFLICT 633 633 R -> I (IN REF. 3).
FT CONFLICT 704 704 A -> R (IN REF. 3).
SQ SEQUENCE 861 AA; 100017 MW; EDD04907BDC9207D CRC64;

Query Match 60.7% Score 34; DB 1; Length 861;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Cy      1 KALVOAMEBLK 12
Db      265 KALINNEQLN 276

RESULT 36
IFH1_YEAST STANDARD; PRT; 1085 AA.
ID IFH1_YEAST
AC P39520;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE IFH1 protein (RRP3 protein).
GN IFH1 OR RRP3 OR YLR223C OR L8083.9.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=FL100;
RX MEDLINE=95304839; PubMed=7785326;
RA Cherel I., Thuriaux P.;
RT "The IFH1 gene product interacts with a fork head protein in
RL Saccharomyces cerevisiae."
RN Yeast 11:261-270(1995).
[2]
SEQUENCE FROM N.A.
RP STRAIN=S288c / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnson M., Hillier L., Riles L., Albertmann K., Andre B., Ansoorge W.,
RA Benes V., Brueckner M., Delfus H., Dubois E., Duesterhoeft A.,
RA Entian K.-D., Floeth M., Goffeau A., Hedling U., Heumann K.,
RA Heuss-Nietzel D., Hilbert H., Hilger F., Kleine K., Koelter P.,
RA Louis E.-J., Messenguy F., Mewes H.-W., Miesga T., Moestl D.,
RA Mueller-Auer S., Nentwich U., Obermayer B., Piravandi E., Pohl T.M.,
RA Portetle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Uristarazu L.A., Vanderbol M., Verhasselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hohnselt J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII."
RN Nature 387:87-90(1997).
CC -1- FUNCTION: CONTROLS THE PRE-RRNA PROCESSING MACHINERY IN
CC CONJUNCTION WITH FHL1. COULD CONVERT FHL1 FROM A REPRESSOR
CC TO AN ACTIVATOR.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z29488; CA82624.1; -
DR EMBL; U19027; AA67412.1; -
DR PIR; S5352; S5352.
DR SGD; S0004213; IFH1.
DR GO; GO:0006348; P:chromatin silencing at telomere; IMP.
KW

```



```

DR GO; GO:0006364; P:RNA processing; IGI.
KM Nucleic protein; Transcription regulation.
FT DOMAIN 122 163 ASP/GLU-RICH (HIGHLY ACIDIC).
SQ SEQUENCE 1085 AA; 122491 MW; BEICDEF06213FE0 CRC64;

Query Match 60.7%; Score 34; DB 1; Length 1085;
Best Local Similarity 54.5%; Pred. No. 2.6e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ALVQOMEDLRQ 12
DB 740 ALVQHQDLRE 750

RESULT 37
AG13 ARATH STANDARD; PRT; 244 AA.
ID _AG13 ARATH STANDARD; PRT; 244 AA.
AC Q38837; Q9W2F0;
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Agamous-like MADS box protein AG13.
CN AG13 OR AF3G61120 OR T20K12.20.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucots II; Brassicales; Brassicaceae; Arabidopsie.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Landsberg erecta; TISSUE=flower;
RX MEDLINE=96004530; PubMed=7549482;
RA Roumlsly S.D.; Ditta G.S.; Yanofsky M.F.;
RT "Diverse roles for MADS box genes in Arabidopsis development.";
RU Plant Cell 7:1259-1269(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=1130713;
RA Salanoubat M.; Lemcke K.; Rieger M.; Ansgore W.; Unselid M.;
RA Fairman B.; Valle G.; Bloeker H.; Perez-Alonso M.; Obermaier B.;
RA Delany M.; Boutry M.; Griwell L.A.; Macho R.; Puigdomenech F.;
RA De Simone V.; Choiane N.; Artiguenave F.; Robert C.; Brothier P.;
RA Winkler P.; Cattoico L.; Weissbach J.; Saurin W.; Quetier F.;
RA Schaefer M.; Mueller-Auer S.; Gabel C.; Fuchs M.; Benes V.;
RA Wurmbach E.; Drzonek H.; Erfle H.; Jordan N.; Bangert S.;
RA Wiedemann R.; Kranz H.; Voss H.; Holland R.; Brandt P.; Nyakatura G.;
RA Verzi A.; D'Angelo M.; Pallavicini A.; Toppo S.; Simionati B.;
RA Conrad A.; Hornischer K.; Kauer G.; Loehner T.-H.; Nordstiek G.;
RA Reichelt J.; Scharfe M.; Schoen O.; Bagues M.; Terol J.; Clement J.;
RA Navarro P.; Collado C.; Perez-Perez A.; Othenwaelder B.; Duchemin D.;
RA Cooke R.; Lande M.; Berger-Llauró C.; Purnelle B.; Maury D.;
RA de Haan M.; Maaree A.C.; Alcaraz J.-P.; Cortez A.; Caescuerra E.;
RA Montfort A.; Argitrou A.; Flores M.; Liguori R.; Vitale D.;
RA Mainhardt G.; Haase D.; Schoof H.; Rudd S.; Zaccaria P.; Mewes H.-W.;
RA Mayor K.F.X.; Kaul S.; Town C.D.; Koo H.L.; Tallon L.J.; Jenkins J.;
RA Rooney T.; Rizzo M.; Walts A.; Utecherback T.; Fujii C.Y.; Shea T.P.;
RA Cressey T.H.; Haas B.; Malt R.; Wu D.; Peterson J.; Van Aken S.;
RA Pat G.; Miltecher J.; Sellers P.; Gill J.E.; Feldlyum T.V.;
RA Preuss D.; Lin X.; Nierman W.C.; Salzberg S.L.; White O.; Venter J.C.;
RA Frazer C.M.; Kaneo T.; Nakamura Y.; Sato S.; Kato T.; Asamizu E.;
RA Sasamoto S.; Kimura T.; Ideasa K.; Kawashima K.; Kishida Y.;
RA Kiyokawa C.; Kohara M.; Matsumoto M.; Matsuno A.; Muraki A.;
RA Nakayama S.; Nakazaki N.; Shino S.; Takeuchi C.; Wada T.;
RA Matenabe A.; Yamada M.; Yasuda M.; Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
thaliana.";
RL Nature 408:820-822(2000).
CC -1- FUNCTION: PROBABLE TRANSCRIPTION FACTOR.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
FACTORS.
CC -1- SIMILARITY: Contains 1 K-box dimerization domain.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U20183; AAC49081.1; -.
DR EMBL; AL137898; CAB71042.1; -.
DR PIR; T47904; T47904.
DR HSSP; P11746; IMNM.
DR TRANSFAC; T03011; -.
DR InterPro; IPR002487; TF_Kbox.
DR InterPro; IPR002100; TF_MADSbox.
DR Pfam; PF01486; K-box; 1.
DR Pfam; PF00319; SRP-TP; 1.
DR PRINTS; PR00404; MADSDOMAIN.
DR SMART; SM00432; MADS; 1.
DR PROSITE; PS00350; MADS_BOX_1; 1.
DR PROSITE; PS50066; MADS_BOX_2; 1.
KM Transcription regulation; DNA-binding; Nuclear protein.
FT DOMAIN 3 57 MADS.
FT DOMAIN 94 166 K-BOX.
FT CONFLICT 213 244 LQIGQHYEGEGSVTKSNARSDAETNFVQ -> YRLGF
FT CONFLICT 213 244 NNTMSKVQDLR (IN REF. 1).
SQ SEQUENCE 244 AA; 27967 MW; 7480EF3F5BBA1499 CRC64;

OY 1 KALVQOMEDLRQ 12
DB 142 QVMEQMEELR 153

RESULT 38
TRPA_BACST STANDARD; PRT; 269 AA.
ID _TRPA_BACST STANDARD; PRT; 269 AA.
AC P19667;
DT 01-FEB-1991 (Rel. 17; Created)
DT 01-FEB-1991 (Rel. 17; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Tryptophan synthase alpha chain (EC 4.2.1.20).
CN TRPA.
OC Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 13737;
RA Ishiwa K.I.; Yoshino S.; Iwamoto S.; Suzuki T.; Makiguchi N.;
RT "Cloning and sequencing of Bacillus stearothermophilus tryptophan
synthase genes.";
RL Agric. Biol. Chem. 53:2941-2948(1989).
CC -1- FUNCTION: The alpha subunit is responsible for the aldol cleavage
of indoleglycerol phosphate to indole and glyceraldehyde 3-
phosphate.
CC -1- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate
= L-tryptophan + glyceraldehyde 3-phosphate.
CC -1- PATHWAY: Tryptophan biosynthesis; fifth (last) step.
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (By
similarity).
CC -1- SIMILARITY: BELONGS TO THE TRPA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/

```



```

CC or send an email to: license@isb-sib.ch).
CC -----
DR EMBL: D00539; BAA00428.1; -.
DR PIR: J050525; J050525.
DR HSSP: P00929; 2MSY.
DR HAMAP: MF_00131; -.
DR InterPro: IPR003009; FMN_enzyme.
DR InterPro: IPR002028; TRP_synthase.
DR Pfam: PF00290; trp_synth. 1.
DR Prodom: PD001535; Trp_synthase; 1.
DR TIGRFAMs: TIGR00262; trpA; 1.
DR PROSITE: PS00167; TRP_SYNTHASE_ALPHA; 1.
DR KMW tryptophan biosynthesis; Lyase.
SQ SEQUENCE 269 AA; 28736 MW; 75D71BA390E1198D CRC64;

Query Match 58.9%; Score 33; DB 1; Length 269;
Best Local Similarity 77.8%; Pred. No. 99;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 ALVQOMQL 10
DB 226 ALVQKVEQL 234

RESULT 39
SRS4_PYRFU STANDARD; PRT; 443 AA.
ID OSU070; O97700;
AC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Signal recognition 54 kDa protein (SRP54).
GN SRP54 OR PF1731.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RA Maeshima H., Okuno E., Aimi T., Morinaga T., Itoh T.;
RT "An archaeal protein homologous to mammalian SRP54 and bacterial Fth
RL submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=VC1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Binds to the signal sequence of presecretory protein
CC when they emerge from the ribosomes (By similarity).
CC -1- SUBUNIT: Archaeal signal recognition particle consists of a 7S RNA
CC molecule of 300 nucleotides and two protein subunits: SRP54 and
CC SRP19 (By similarity).
CC -1- SUCCESFULLAR LOCATION: Cytoplasmic (By similarity).
CC -1- DOMAIN: Has a two domain structure: the G-domain binds GTP; the M-
CC domain binds the 7S RNA in presence of SRP19 and also binds the M-
CC signal sequence (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SRP FAMILY OF GTP-BINDING PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB057373; BAB64926.1; -.
DR EMBL: AB010270; AAL81855.1; -.
DR HAMAP: MF_00306; -.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR000897; SRP54.

```

```

DR InterPro: IPR004125; SRP54_SPB.
DR Pfam: PF00448; SRP54_N; 1.
DR Pfam: PF02881; SRP54_N; 1.
DR Pfam: PF02978; SRP_SPB; 1.
DR Prodom: PD000819; SRP54; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00300; SRP54; 1.
DE Signal recognition particle; GTP-binding; RNA-binding;
KW Complete proteome.
FT DOMAIN 1 294 G-DOMAIN (BY SIMILARITY).
FT DOMAIN 295 443 M-DOMAIN (BY SIMILARITY).
FT NP BIND 107 114 GTP (BY SIMILARITY).
FT NP BIND 189 193 GTP (BY SIMILARITY).
FT NP BIND 247 250 GTP (BY SIMILARITY).
FT CONFLICT 150 150 L -> V (IN REF. 1).
SQ SEQUENCE 448 AA; 49879 MW; 102D1717B207F71A CRC64;

Query Match 58.9%; Score 33; DB 1; Length 443;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 1 KALVQOMQL 10
DB 198 KALIEEMKQI 207

RESULT 40
RMUC_ZYMO STANDARD; PRT; 448 AA.
ID RMUC_ZYMO
AC Q9REQ3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA recombination protein rmuc homolog.
GN RMUC.
OS Zymomonas mobilis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Zymomonas.
OX NCBI_TaxID=542;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 31821 / ZM4 / CP4;
RA Um H.W., Kang H.S.;
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Involved in DNA recombination (By similarity).
CC -1- SIMILARITY: BELONGS TO THE RMUC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF213832; AAF23792.1; -.
DR InterPro: IPR003798; DUF195.
DR Pfam: PF02646; Rmuc; 1.
KW DNA recombination; Coiled coil.
FT DOMAIN 43 167 COILED COIL (POTENTIAL).
SQ SEQUENCE 448 AA; 50936 MW; AA9F65D3E2182E5E CRC64;

Query Match 58.9%; Score 33; DB 1; Length 448;
Best Local Similarity 58.3%; Pred. No. 1.6e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 KALVQOMQL 12
DB 154 EASYKAEQL 165

Search completed: September 8, 2003, 16:05:24
Job time : 12 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 8, 2003, 16:03:12 ; Search time 34 Seconds  
(without alignments)  
91.077 Million cell updates/sec

Title: US-09-991-809-1  
Perfect score: 56  
Sequence: 1 KALVQMEQLRQ 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_prodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	244	4 Q13784	Q13784 homo sapien
2	51	91.1	91	6 Q28857	Q28857 pan troglod
3	49	87.5	395	11 Q9DBNO	Q9DBNO mus musculu
4	49	87.5	395	11 Q91XR8	Q91XR8 mus musculu
5	49	87.5	435	11 Q01488	Q01488 mus musculu
6	41	73.2	398	11 Q8R0V5	Q8R0V5 mus musculu
7	41	73.2	613	2 Q9F7L4	Q9F7L4 uncultured
8	40	71.4	313	13 Q9PUA6	Q9PUA6 xenopus lae
9	40	71.4	360	2 Q8G1Q8	Q8G1Q8 aeromonas s
10	39	69.6	311	13 Q90370	Q90370 coturnix co
11	39	69.6	311	13 Q90888	Q90888 gallus gall
12	39	69.6	460	11 Q8C459	Q8C459 mus musculu
13	39	69.6	2712	10 Q9SB74	Q9SB74 arabidopsis
14	38	67.9	259	2 Q93PU6	Q93PU6 pseudomonas
15	38	67.9	259	2 Q9R9U0	Q9R9U0 pseudomonas
16	38	67.9	498	4 Q9BT92	Q9BT92 homo sapien

17	38	67.9	498	4 Q8NAG0	Q8NAG0 homo sapien
18	38	67.9	830	16 Q8XR09	Q8XR09 ralsiconia s
19	37	66.1	177	17 Q28120	Q28120 archaeoglob
20	37	66.1	254	11 Q9DA73	Q9DA73 mus musculu
21	37	66.1	359	16 Q8VNO7	Q8VNO7 anabaena sp
22	37	66.1	385	16 Q8DHP3	Q8DHP3 synechococ
23	37	66.1	480	16 Q8G4H0	Q8G4H0 bifidobacte
24	37	66.1	490	16 Q8F4J0	Q8F4J0 leptospira
25	37	66.1	893	16 Q8YWM9	Q8YWM9 anabaena sp
26	37	66.1	900	17 Q8PUB5	Q8PUB5 methanosaic
27	36	64.3	126	12 Q9EMW9	Q9EMW9 amasacta moo
28	36	64.3	196	4 Q9H735	Q9H735 homo sapien
29	36	64.3	208	16 Q9RX92	Q9RX92 deinococcus
30	36	64.3	264	4 Q8N2T9	Q8N2T9 homo sapien
31	36	64.3	354	11 Q9CUG5	Q9CUG5 mus musculu
32	36	64.3	362	2 Q9AES7	Q9AES7 aeromonas h
33	36	64.3	370	16 Q8ZSL2	Q8ZSL2 salmonella
34	36	64.3	370	16 Q9XDN0	Q9XDN0 salmonella
35	36	64.3	380	5 Q9V4Y9	Q9V4Y9 drosophila
36	36	64.3	498	16 Q9KBS6	Q9KBS6 bacillus ha
37	36	64.3	521	5 Q8MSA0	Q8MSA0 drosophila
38	36	64.3	530	4 Q9BWC1	Q9BWC1 homo sapien
39	36	64.3	577	5 Q9VDS2	Q9VDS2 drosophila
40	36	64.3	674	4 Q9H7J2	Q9H7J2 homo sapien
41	36	64.3	733	2 Q9S6G9	Q9S6G9 helicobacte
42	36	64.3	778	16 Q9FBP3	Q9FBP3 streptomyce
43	36	64.3	965	11 Q8C759	Q8C759 mus musculu
44	36	64.3	1279	5 Q46099	Q46099 drosophila
45	36	64.3	1589	5 Q8MY08	Q8MY08 dictyosteli

## ALIGNMENTS

## RESULT 1

Q13784 PRELIMINARY; PRT; 244 AA.  
ID Q13784;  
AC Q13784;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE APOA4 protein (Fragment).  
GN APOA4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=86016704; PubMed=3931073;  
RA Karathanasis S.K.;  
RT "Apolipoprotein multigene family: tandem organization of human  
RT Apolipoprotein AI, CIII, and AIV genes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:6374-6378(1985).  
DR EMBL: M10373; AAB59516.1; -.  
DR HSSP; P02649; 1BZ4.  
DR InterPro; IPR00074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein.1.  
FT NON TER 1  
SQ SEQUENCE 244 AA; 28157 MW; 01616ACB53F12734 CRC64;

Query Match 100.0%; Score 56; DB 4; Length 244;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVQMEQLRQ 12  
|||  
Db 182 KALVQMEQLRQ 193

RESULT 2  
Q28857

```

ID Q2857 PRELIMINARY; PRT; 91 AA.
AC Q2857;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Apolipoprotein A-IV (Fragment).
GN APOA4.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94374966; PubMed=8088751;
RA Kamboh M.I., Kelly L.J., Ann Y.I., Ferrell R.E.;
RT "Genetic polymorphism of apolipoprotein A-IV in the chimpanzee: common
RL deletion of a conserved 12-nucleotide tandem repeat.";
DR Hum. Biol. 66:625-638(1994).
KW EMBL; 572300; AAB31973.1; -.
GN Lipoprotein.
FT NON TER
SQ SEQUENCE 91 AA; 10619 MW; 5D654D84D3A4A0 CRC64;

Query Match 91.1%; Score 51; DB 6; Length 91;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVQOMEOLR 11
DB 11 KALVQOMEOLR 21

RESULT 3
Q9DBNO PRELIMINARY; PRT; 395 AA.
ID Q9DBNO;
AC Q9DBNO;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Adult male liver cDNA, RIKEN full-length enriched library,
DE clone:1300002K10, full insert sequence.
GN APOA4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=21085660; PubMed=11217651;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stabili F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Caminici P., de Bonaldo M.F.,
RA Brownstein W.J., Bult C., Fletcher N., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Maehama M., Mazzarelli J., Mombaerts P.,
RA Norone P., Ring B., Ringwald C., Rodriguez I., Sakamoto N.,
RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK004856; BAB23620.1; -.
DR MGD; MGI:88051; Apoa4.
DR InterPro; IPR000074; Apolipoprotein.

```

```

DR Pfam; PF01442; Apolipoprotein; 2.
SQ SEQUENCE 395 AA; 45044 MW; 4102B84ACB0D182A CRC64;

Query Match 87.5%; Score 49; DB 11; Length 395;
Best Local Similarity 83.3%; Pred. No. 0.94;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KALVQOMEOLRQ 12
DB 316 KALVQOLEQFRQ 327

RESULT 4
Q91XF8 PRELIMINARY; PRT; 395 AA.
ID Q91XF8;
AC Q91XF8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Unknown (Protein for MGC:18592).
GN APOA4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Kidney;
RX TISSUE=Kidney;
RA Strassberg R.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC010769; AAH10769.1; -.
DR MGD; MGI:88051; Apoa4.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 2.
SQ SEQUENCE 395 AA; 45029 MW; C48BE32EBD441F71 CRC64;

Query Match 87.5%; Score 49; DB 11; Length 395;
Best Local Similarity 83.3%; Pred. No. 0.94;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KALVQOMEOLRQ 12
DB 316 KALVQOLEQFRQ 327

RESULT 5
Q01488 PRELIMINARY; PRT; 435 AA.
ID Q01488;
AC Q01488;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Apolipoprotein A-IV precursor.
GN APOA4.
OS Mus musculus castaneus (southeastern Asian house mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10091;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CAST/EiJ; TISSUE=Liver;
RX MEDLINE=91286309; PubMed=1648102;
RA Reue K., Leete T.H.;
RT "Genetic variation in mouse apolipoprotein A-IV due to insertion and
RT deletion in a region of tandem repeats.";
RL J. Biol. Chem. 266:12715-12721(1991).
DR EMBL; M64250; AAA37216.1; -.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 2.
KW Signal; Lipoprotein.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 435 APOLIPOPROTEIN A-IV.
SQ SEQUENCE 435 AA; 49254 MW; B2AF55BEA8E456B2 CRC64;

```

Query Match 87.5%; Score 49; DB 11; Length 435;  
Best Local Similarity 83.3%; Pred. No. 1;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12  
|||:|||||  
DB 316 KALVOQLEQFRQ 327

## RESULT 6

Q8R0V5 PRELIMINARY; PRT; 398 AA.  
AC Q8R0V5;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
OS Hypothetical 44.4 kDa protein.  
OC Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Strausberg R.;  
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC026393; AA026393.1; -  
DR InterPro: IPR000898; IDO\_fam.  
DR Pfam: PF01231; IDO; 1.  
DR Hypothetical protein.  
SQ SEQUENCE 398 AA; 44439 MW; 7871CC86FE24D1A CRC64;

Query Match 73.2%; Score 41; DB 11; Length 398;  
Best Local Similarity 75.0%; Pred. No. 26;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12  
|||:|||||  
DB 176 KALVOQMEAIRQ 187

## RESULT 7

Q9F7L4 PRELIMINARY; PRT; 613 AA.  
AC Q9F7L4;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE Predicted preprotein translocase secA subunit (Fragment).  
OS uncultured proteobacterium EBAC31A08.  
OC Bacteria; Proteobacteria; environmental samples.  
OX NCBI\_TaxID=133804;  
RN [1]  
RP SEQUENCE FROM N.A.  
RP MEDLINE=20446260; PubMed=10988064;  
RA Beja O., Aravind L., Koonin E.V., Suzuki M.T., Hadd A., Nguyen L.P.,  
RA Jovanovich S.B., Gates C.M., Feldman R.A., Spudich J.L., Spudich E.N.,  
RA "Bacterial rhodopsin: evidence for a new type of phototrophy in the  
RT Sea.";  
RL Science 289:1902-1906 (2000).  
DR EMBL: AF279106; AAG10505.1; -  
DR InterPro: IPR000185; SecA.  
DR Pfam: PF01043; SecA\_protein; 1.  
DR PRINTS: PRO0906; SECA.  
DR PROSITE: PS01312; SECA; 1.  
FT NON TER  
SQ SEQUENCE 613 AA; 70109 MW; FFF95DB3C96BD1C3 CRC64;

Query Match 73.2%; Score 41; DB 2; Length 613;  
Best Local Similarity 58.3%; Pred. No. 39;  
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12  
|||:|||||  
DB 148 KALIEIEQLRK 159

## RESULT 8

Q9PUA6 PRELIMINARY; PRT; 313 AA.  
AC Q9PUA6;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE bZIP transcription factor mafB.  
GN MAFB.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21152895; PubMed=11231068;  
RA Ishibashi S., Yasuda K.;  
RT "Distinct roles of maf genes during Xenopus lens development.";  
RL Mech. Dev. 101:155-166 (2001).  
DR EMBL: AF202058; AAF08316.1; -  
DR InterPro: IPR004827; TP\_bZIP.  
DR InterPro: IPR004826; TP\_Maf.  
DR Pfam: PF03131; bZIP\_Maf; 1.  
DR SMART: SM00338; BRLT; 1.  
SQ SEQUENCE 313 AA; 35714 MW; 8E697A0A928BF95 CRC64;

Query Match 71.4%; Score 40; DB 13; Length 313;  
Best Local Similarity 80.0%; Pred. No. 31;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LVQOQMEQLRQ 12  
|||:|||||  
DB 260 LVQOQVEQLKQ 269

## RESULT 9

Q8GLQ8 PRELIMINARY; PRT; 360 AA.  
AC Q8GLQ8;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE lafe.  
GN lafe.  
OS Aeromonas salmonicida.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Aeromonadales;  
OC Aeromonadaceae; Aeromonas.  
OX NCBI\_TaxID=645;  
RN [1]  
RP SEQUENCE FROM N.A.  
RP Tomas J.M., Merino S.;  
RA "A colonization factor (production of lateral flagella) of mesophilic  
RT Aeromonas is inactive in all A. salmonicida strains.";  
RL Submitted (Jul-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY129557; AAN04221.1; -  
SQ SEQUENCE 360 AA; 39551 MW; C6128656AC46431A CRC64;

Query Match 71.4%; Score 40; DB 2; Length 360;  
Best Local Similarity 80.0%; Pred. No. 36;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALVOQMEQLR 11  
|||:|||||  
DB 293 ALIOMSEQLR 302

## RESULT 10

Q90370 PRELIMINARY; PRT; 311 AA.  
 AC Q90370;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE MAF protein.  
 GN MAF.  
 OS Coturnix coturnix japonica (Japanese quail).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Coturnix.  
 RX NCBI\_TaxID=93934;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96180718; PubMed=8620536;  
 RA Sieweke M.H., Tekotte H., Frampton J., Graf T.,  
 RT "MAF is an interaction partner and repressor of Ets-1 that inhibits  
 erythroid differentiation."  
 RL Cell 85:49-60(1996).  
 DR EMBL: X96511; CAA65360.1; -;  
 DR InterPro: IPR004827; TF\_bZIP.  
 DR InterPro: IPR004826; TF\_Maf.  
 DR Pfam: PF03131; bZIP\_Maf; 1.  
 DR SMART: SM00338; BRLZ; 1.  
 SQ SEQUENCE 311 AA; 35476 MW; 7D1F3FA05D5CD683 CRC64;

Query Match 69.6%; Score 39; DB 13; Length 311;  
 Best Local Similarity 70.0%; Pred. No. 47;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LVQOMEQLRQ 12  
 |::|::|::|  
 DB 261 LIQVEQLKQ 270

RESULT 11  
 Q90888 PRELIMINARY; PRT; 311 AA.  
 ID Q90888;  
 AC Q90888;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE MAF.  
 GN Gallus gallus (Chicken).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 RX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95021288; PubMed=7935473;  
 RA Katoka K., Fujiwara K.T., Noda M., Nishizawa M.;  
 RT "MAF, a new Maf family transcription activator that can associate with  
 Maf and Fos but not with Jun."  
 RL Mol. Cell. Biol. 14:7581-7591(1994).  
 DR EMBL: D28600; BAA05938.1; -;  
 DR InterPro: IPR004827; TF\_bZIP.  
 DR InterPro: IPR004826; TF\_Maf.  
 DR Pfam: PF03131; bZIP\_Maf; 1.  
 DR SMART: SM00338; BRLZ; 1.  
 SQ SEQUENCE 311 AA; 35467 MW; DDAE7F69887D3ABA CRC64;

Query Match 69.6%; Score 39; DB 13; Length 311;  
 Best Local Similarity 70.0%; Pred. No. 47;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LVQOMEQLRQ 12  
 |::|::|::|  
 DB 261 LIQVEQLKQ 270

## RESULT 12

Q8C459 PRELIMINARY; PRT; 460 AA.  
 ID Q8C459;  
 AC Q8C459;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN Mus musculus (Mouse).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 DR EMBL: AK083027; BAC38735.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 460 AA; 55949 MW; FBE7110A51BE37B9 CRC64;

Query Match 69.6%; Score 39; DB 11; Length 460;  
 Best Local Similarity 66.7%; Pred. No. 68;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KALVQMEQLRQ 12  
 |::|::|::|  
 DB 217 EALRQMEQLKQ 228

RESULT 13  
 Q9SB74 PRELIMINARY; PRT; 2712 AA.  
 ID Q9SB74;  
 AC Q9SB74;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical 304.1 kDa protein.  
 GN F28M20.240 OR A7G31570.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 RX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Rieger M., Mueller-Auer S., Zipp M., Schaefer M.,  
 RA Hohelsel J., Neues H.W., Mayer K.F.X., Schueller C.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Murphy G., Ridley P., Hudson S., Neues H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Rieger M., Mueller-Auer S., Zipp M., Schaefer M., Neues H.W.,  
 RA Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL031004; CAA19766.1; -;  
 DR EMBL: AL161579; CAB79875.1; -;  
 KW Hypothetical protein.

SO SEQUENCE 2712 AA; 304052 MW; 75ADBE60C4891ACF CRC64;  
Query Match 69.6%; Score 39; DB 10; Length 2712;  
Best Local Similarity 66.7%; Pred. No. 3.7e+02;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 KALVOQMEQLRQ 12  
Db 1873 KALVQORDSLKQ 1884  
RESULT 14  
Q93PU6 PRELIMINARY; PRT; 259 AA.  
AC Q93PU6  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Putative regulatory protein ItgV.  
GN ITGV.  
OS Pseudomonas putida.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DOT-T1E;  
RX MEDLINE=21289082; PubMed=11395460;  
RA Rojas A., Dugue E., Mosqueda G., Golden G., Hurtado A., Ramos J.L.,  
Segura A.,  
RT "Three Efflux Pumps Are Required To Provide Efficient Tolerance to  
Toluene in Pseudomonas putida DOT-T1E."  
RL J. Bacteriol. 183:3967-3973 (2001).  
CC -1- SIMILARITY: BELONGS TO THE ICLR FAMILY OF TRANSCRIPTIONAL  
CC REGULATORS.  
DR EMBL; AF299253; AAK69562.1;  
DR InterPro; IPR005471; HTH\_ICLR.  
DR Pfam; PF01614; ICLR.1.  
DR SMART; SM00346; HTH\_ICLR.1.  
KW DNA-binding; Transcription; Transcription regulation.  
SQ SEQUENCE 259 AA; 27410 MW; 1B2A580D9A85767 CRC64;  
QY Query Match 67.9%; Score 38; DB 2; Length 259;  
Best Local Similarity 58.3%; Pred. No. 60;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
Db 179 KALVQOLSEVRQ 190  
RESULT 15  
Q9R9U0 PRELIMINARY; PRT; 259 AA.  
AC Q9R9U0  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE Efflux pump regulator SrpS.  
GN SRPS.  
OS Pseudomonas putida.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SL2;  
RA Dennis J.J., Zylstra G.J.;  
RT "Characterization of the Regulators of the Srp Solvent Resistance  
Efflux Pump."  
RL Submitted (APR-1998) to the EMBL/Genbank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE ICLR FAMILY OF TRANSCRIPTIONAL  
CC REGULATORS.

DR EMBL; AF061937; AAF16681.1; -;  
DR HSSP; P03020; 2CGP.  
DR InterPro; IPR005471; HTH\_ICLR.  
DR Pfam; PF01614; ICLR.1.  
DR SMART; SM00346; HTH\_ICLR.1.  
KW DNA-binding; Transcription; Transcription regulation.  
SQ SEQUENCE 259 AA; 27579 MW; 322E1DF03842C211 CRC64;  
QY Query Match 67.9%; Score 38; DB 2; Length 259;  
Best Local Similarity 58.3%; Pred. No. 60;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
Db 179 KALVQOLSEVRQ 190  
QY 1 KALVOQMEQLRQ 12  
Db 179 KALVQOLSEVRQ 190  
RESULT 16  
Q9BT92 PRELIMINARY; PRT; 498 AA.  
AC Q9BT92  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE Similar to RIKEN cDNA A930031F18 gene.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RA Strauberg R.;  
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.  
DR EMBL; BC004285; AA04285.1; -;  
SQ SEQUENCE 498 AA; 61071 MW; 7D1E8A66C56516F8 CRC64;  
QY Query Match 67.9%; Score 38; DB 4; Length 498;  
Best Local Similarity 63.6%; Pred. No. 1.1e+02;  
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
Db 218 EALVQMEELK 228  
RESULT 17  
Q8NAGO PRELIMINARY; PRT; 498 AA.  
AC Q8NAGO  
DT 01-OCT-2002 (TREMBlrel. 22, Created)  
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Hypothetical protein FLJ35417.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Small intestine;  
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,  
Horuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J., Chiba Y.,  
Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,  
Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,  
Yamamoto J., Isono Y., Kawai-Hio Y., Sato K., Nishikawa T.,  
Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,  
Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,  
Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K., Isogai T.,  
RA "NEO human cDNA sequencing project."  
RT Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AK092736; BAC03960.1; -;  
KW Hypothetical protein.

```

SQ SEQUENCE 498 AA; 61099 MW; DCCDC35B793869A3 CRC64;
Query Match
Best Local Similarity 67.9%; Score 38; DB 4; Length 498;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 KALVOOMEOLRQ 11
    |||:|||||:
DB 218 BALLOOMEELK 228

RESULT 18
O8XR09 PRELIMINARY; PRT; 830 AA.
AC O8XR09;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Putative transmembrane protein.
GN RSP0732 OR RS05796.
OS Plasmodium megaloschizont.
OC Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Ralstoniaceae; Ralstonia.
OX NCBI_TaxId=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Caticolico L.,
RA Chandelier M., Choise N., Claudel-Renard C., Cumac S., Demange N.,
RA Gagnier C., Lave M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigrier P., Trebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
RL Nature 415:497-502(2002).
DR EMBL; AL646080; CAD17883.1; -.
KW Plasmid; Complete proteome.
SQ SEQUENCE 830 AA; 89945 MW; A13D0A4782C865D7 CRC64;

Query Match
Best Local Similarity 67.9%; Score 38; DB 16; Length 830;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KALVOOMEOLRQ 12
    |||:|||||:
DB 543 RALVPOREOLR 554

RESULT 19
O28120 PRELIMINARY; PRT; 177 AA.
AC O28120;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-MAR-2002 (TREMBLrel. 05, Last sequence update)
DE Hypothetical protein AF2162.
GN AF2162.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxId=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirtrees E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,

```

```

RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spirig T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL; AE000955; AAB89098.1; -.
DR TIGR; AF2162; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 177 AA; 20384 MW; EDD4211EB05FD75D CRC64;

Query Match
Best Local Similarity 66.1%; Score 37; DB 17; Length 177;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 KALVOOMEOLRQ 12
    |||:|||||:
DB 153 KALIPOLDLRQ 164

RESULT 20
O9DA73 PRELIMINARY; PRT; 254 AA.
AC O9DA73;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 1700019B01Rik protein.
GN 1700019B01Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boileau D., Bolunga N., Carninci P., de Bona M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Guenincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombae P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta Y., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlschki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK006105; BMB24410.1; -.
DR MGD; MGI:1917304; 1700019B01Rik.
SQ SEQUENCE 254 AA; 29644 MW; 82422377492A5025 CRC64;

Query Match
Best Local Similarity 66.1%; Score 37; DB 11; Length 254;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 KALVOOMEOLRQ 12
    |||:|||||:
DB 64 KALVELEVLKQ 75

RESULT 21

```



Q8YN07  
ID Q8YN07 PRELIMINARY; PRT; 359 AA.  
AC Q8YN07;  
DT 01-MAR-2002 (TRENBLrel. 20, Created)  
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
DE Hypothetical protein ALR4505.  
GN ALR4505.  
OS Anaerobaculum sp. (strain PCC 7120).  
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
OX NCBI\_TaxID=103690;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21595285; PubMed=11759840;  
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Moraki A.,  
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,  
RA Yasuda M., Tabata S.;  
RT "Complete genomic sequence of the filamentous nitrogen-fixing  
RT cyanobacterium Anaerobaculum sp. strain PCC 7120.";  
RL DNA Res. 8:205-213(2001).  
DR EMBL; AP003596; BMB76204.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 359 AA; 41569 MW; 2E839DCAE10F0369 CRC64;

Query Match 66.1%; Score 37; DB 16; Length 359;  
Best Local Similarity 58.3%; Pred. No. 1.2e+02;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12  
|||:|:|:|:  
Db 210 KALYQGLELQ 221

RESULT 22  
Q8DHP3  
ID Q8DHP3 PRELIMINARY; PRT; 385 AA.  
AC Q8DHP3;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE TLR1902 protein.  
GN TLR1902.  
OS Synecococcus elongatus (Thermosynechococcus elongatus).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.  
OX NCBI\_TaxID=32046;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2225144; PubMed=12240834;  
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,  
RA Watanabe A., Iriyuchi M., Kawashima K., Kimura T., Kishida Y.,  
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,  
RA Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the thermophilic cyanobacterium  
RT Thermosynechococcus elongatus BP-1.";  
RL DNA Res. 9:123-130(2002).  
DR EMBL; AP003375; BAC09454.1; -;  
KW Complete proteome.  
SQ SEQUENCE 385 AA; 43975 MW; 85DF3980FAF29B76 CRC64;

Query Match 66.1%; Score 37; DB 16; Length 385;  
Best Local Similarity 72.7%; Pred. No. 1.3e+02;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALVQOMELRQ 12  
|||:|:|:|:  
Db 5 ALSQKIEQLRQ 15

RESULT 23  
Q8G4H4

ID Q8G4H4 PRELIMINARY; PRT; 480 AA.  
AC Q8G4H4;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Widely conserved hypothetical protein in upl0004.  
GN BL1408.  
OS Bifidobacterium longum.  
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;  
OC Bifidobacteriaceae; Bifidobacterium.  
OX NCBI\_TaxID=216816;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=NCC 2705;  
RA MEDLINE=22294977; PubMed=12381787;  
RA Schell M.A., Karmaliantzou M., Snel B., Vlianova D., Berger B.,  
RA Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,  
RA Prilmore R.D., Atigoni F.;  
RT "The genome sequence of Bifidobacterium longum reflects its adaptation  
RT to the human gastrointestinal tract.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).  
DR EMBL; AE014771; AAN25207.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 480 AA; 52671 MW; 2B1D327943FE9BFE CRC64;

Query Match 66.1%; Score 37; DB 16; Length 480;  
Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12  
|||:|:|:|:  
Db 148 KAQVQTEQLRQ 159

RESULT 24  
Q8F4J0  
ID Q8F4J0 PRELIMINARY; PRT; 490 AA.  
AC Q8F4J0;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE UDP-N-acetylmutamate--alanine ligase (EC 6.3.2.8).  
GN MURC OR LA2051.  
OS Leptospira interrogans.  
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
OX NCBI\_TaxID=173;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;  
RA Ren S.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE011377; AAN49250.1; -;  
KW Ligase; Complete proteome.  
SQ SEQUENCE 490 AA; 54828 MW; C29A9FB0C9829635 CRC64;

Query Match 66.1%; Score 37; DB 16; Length 490;  
Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 KALVOQMEQLRQ 12  
|||:|:|:|:  
Db 375 KAVIOSMELCK 386

RESULT 25  
Q8YUJ9  
ID Q8YUJ9 PRELIMINARY; PRT; 893 AA.  
AC Q8YUJ9;  
DT 01-MAR-2002 (TRENBLrel. 20, Created)  
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Toxin secretion ABC transporter ATP-binding protein.  
GN ALR1614.

```

OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxId=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Ref. 8:205-213(2001).
DR EMBL, AP003586; BAB77980.1; -.
DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR001140; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR000595; ABC_binding.
DR Pfam: PF00664; ABC_membrane; 1.
DR Pfam: PF00005; ABC_tran; 1.
DR ProDom: PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00100; CNMP; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR PROSITE; PS00442; CNMP_BINDING; 3; 1.
DR ATP-binding; Complete proteome.
KW KW
SQ SEQUENCE 893 AA; 99819 MW; 3183EB0E50227C70 CRC64;

Query Match 66.1%; Score 37; DB 16; Length 893;
Best Local Similarity 63.6%; Pred. No. 3e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 ALVQOMEQLRQ 12
DB 437 ARVQLEQLRQ 447

RESULT 26
O8PUB5 PRELIMINARY; PRT; 900 AA.
ID O8PUB5;
AC O8PUB5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE DNA gyrase, subunit A (EC 5.99.1.3).
GN MM2420.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxId=2209;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Goel / Goel / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Depierremer U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wierer A., Baenauer S., Jacobi C.,
RA Bruggemann H., Lienard T., Christmann A., Boemcke M., Seckel S.,
RA Bhattacharya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-U., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL; AE013485; AMM32116.1; -.
DR InterPro: IPR005743; DNA_gyrase.
DR InterPro: IPR006691; DNA_gyrase_C.
DR InterPro: IPR002205; DNA_topoisom.
DR Pfam: PF03989; DNA_gyrase_C; 6.
DR Pfam: PF00521; DNA_topoisom; 1.
DR ProDom: PD000742; DNA_topoisom; 1.
DR TIGRFAMs; TIGR01063; gyra; 1.
KW Isoenzyme; Complete proteome.
SQ SEQUENCE 900 AA; 101310 MW; A471287E39124225 CRC64;

```

```

Query Match 66.1%; Score 37; DB 17; Length 900;
Best Local Similarity 58.3%; Pred. No. 3e+02;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 1 KALVQMEQLRQ 12
DB 454 EALVQIEELRK 465

RESULT 27
O9EMW9 PRELIMINARY; PRT; 126 AA.
ID O9EMW9;
AC O9EMW9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE AMV080.
GN AMV080.
OS Amaseta moorei entomopoxvirus (AmEPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
OC Entomopoxvirus B.
OX NCBI_TaxId=28321;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20396580; PubMed=10936094;
RA Bowden A.L., Glaesberg K.J., Digans J., Shaw R., Farmerie W.,
RA Moyer R.W.;
RT "Complete Genomic Sequence of the Amaseta moorei Entomopoxvirus:
RT Analysis and Comparison with Other Poxviruses.";
RL Virology 274:120-139(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Bowden A.L., Glaesberg K.J., Digans J., Shaw R., Farmerie W.,
RA Moyer R.W.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF250284; AAG02786.1; -.
SQ SEQUENCE 126 AA; 15109 MW; C090AEFC8D95EC0 CRC64;

Query Match 64.3%; Score 36; DB 12; Length 126;
Best Local Similarity 60.0%; Pred. No. 69;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 ALVQOMEQLR 11
DB 60 ALVQLEELR 69

RESULT 28
O9H735 PRELIMINARY; PRT; 196 AA.
ID O9H735;
AC O9H735;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ21438.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Colon;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Ohtani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isegai T., Sugano S.;
RT "NEBO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK025091; BAB15064.1; -.
KW Hypothetical protein.
SQ SEQUENCE 196 AA; 22323 MW; B0FE9E3B2A8A0A59 CRC64;

Query Match 64.3%; Score 36; DB 4; Length 196;

```

Best Local Similarity 63.6%; Pred. No. 1.1e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
OY 1 KALVOQMEQLR 11  
: || : || || ||  
Db 173 KALTEQOEQLR 183

RESULT 29  
O9RX92 PRELIMINARY; PRT; 208 AA.  
AC O9RX92.  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
DE Hypothetical protein DR0423.  
GN DR0423.  
OS Deinococcus radiodurans.  
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;  
OC Deinococcaceae; Deinococcus.  
OX NCBI\_TaxID=1299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RI;  
RX MEDLINE=20036896; PubMed=10567266;  
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
RA Vamathevan J.J., Lam P., McDonald L., Uetzelack T., Zalewski C.,  
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
RA Fraser C.M.;  
RT "Genome sequence of the radioresistant bacterium Deinococcus  
RT radiodurans R1."  
RL Science 286:1571-1577(1999).  
DR EMBL; AE001902; AAF10011.1; -.  
DR TIGR; DR0423; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 208 AA; 23002 MW; D91A2C42DA086A53 CRC64;

Query Match 64.3%; Score 36; DB 16; Length 208;  
Best Local Similarity 88.9%; Pred. No. 1.1e+02;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 3 LVQOMEQLR 11  
||| ||||  
Db 157 LVQAMEQLR 165

RESULT 30  
O8N2T9 PRELIMINARY; PRT; 264 AA.  
AC O8N2T9.  
DT 01-OCT-2002 (TRENBLrel. 22, Created)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Hypothetical protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAUSBERG R.;  
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC030281; AAH30281.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 264 AA; 29979 MW; BBE07204F881DE12 CRC64;

Query Match 64.3%; Score 36; DB 4; Length 264;  
Best Local Similarity 63.6%; Pred. No. 1.4e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KALVOQMEQLR 11  
: || : || || ||  
Db 173 KALTEQOEQLR 183

RESULT 31  
O9CUG5 PRELIMINARY; PRT; 354 AA.  
AC O9CUG5.  
DT 01-JUN-2001 (TRENBLrel. 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
DE 4930564N1SR1K protein (Fragment).  
GN 4930564N1SR1K.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai T., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aitawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,  
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schintl L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kontseki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
DR EMBL; AK016222; BAB30155.1; -.  
DR MCD; MGI:1922555; 4930564N1SR1K.  
DR InterPro; IPR002110; ANK.  
DR Pfam; PF00023; ank; 3.  
DR SMART; SMO0248; ANK; 3.  
DR PROSITE; PS50086; ANK\_REPEAT; 2.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
KW ANK repeat; Repeat.  
FT NON TER 354 354  
SQ SEQUENCE 354 AA; 40365 MW; 8DDAE04539C37CF1 CRC64;

Query Match 64.3%; Score 36; DB 11; Length 354;  
Best Local Similarity 63.6%; Pred. No. 1.8e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
OY 1 KALVOQMEQLR 11  
||| ||||  
Db 189 KALQAMEQLR 199

RESULT 32  
O9AES7 PRELIMINARY; PRT; 362 AA.  
AC O9AES7.  
DT 01-JUN-2001 (TRENBLrel. 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE Flagellar hook control length protein.  
GN LAFE.  
OS Aeromonas hydrophila.

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Aeromonadales;
OC Aeromonadaceae; Aeromonas.
OX NCBI_TaxID=644;
RN [1]
RP SEQUENCE FROM N.A.
RA Tomas J.M., Merino S., Gavin R.;
RT "Molecular characterization of the Aeromonas lateral flagella.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY028400; AAK20923.1; -.
SQ SEQUENCE 362 AA; 39643 MW; 95FAA2DC4F3DE112 CRC64;

Query Match          64.3%; Score 36; DB 2; Length 362;
Best Local Similarity 70.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALVQOMEQLRQ 11
DB 295 ALIQSMERLR 304

RESULT 33
0825L2 PRELIMINARY; PRT; 370 AA.
AC 0825L2;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Purative propanol dehydrogenase (EC 1.1.1.-).
GN PDUO OR STY2257.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Churchill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Crofton A., Davis P., Davies R.M., Dowd L., White N., Farrer J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA Krogan A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
DR EMBL; AL627273; CAD02413.1; -.
DR InterPro; IPR001670; Fe-ADH.
DR Pfam; PF00465; Fe-ADH; 1.
DR PROSITE; PS00913; ADH_IRON_1; 1.
KW Oxidoreductase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 370 AA; 39435 MW; 490DF9EB3C9858 CRC64;

Query Match          64.3%; Score 36; DB 16; Length 370;
Best Local Similarity 63.6%; Pred. No. 1.9e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALVQOMEQLRQ 12
DB 305 ALIQOIELLKQ 315

RESULT 34
Q9XDN0 PRELIMINARY; PRT; 370 AA.
AC Q9XDN0;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Propanediol utilization: propanol dehydrogenase (PduO).
GN PDUO OR STM2052.

```

```

OS Salmonella typhimurium, and
OS Salmonella enterica subsp. enterica serovar Typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602, 90371;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.enterica subsp. enterica serovar Typhimurium; STRAIN=LT2;
RX MEDLINE=94350831; PubMed=8071226;
RA Chen P., Andersson D.I., Roth J.R.;
RT "The control region of the pdu/cob regulon in Salmonella
RT typhimurium.";
RL J. Bacteriol. 176:5474-5482(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.enterica subsp. enterica serovar Typhimurium; STRAIN=LT2;
RX MEDLINE=98012959; PubMed=9352910;
RA Bobik T.A., Xu Y., Vetter R.W., Otto K.E., Roth J.R.;
RT "Propanediol utilization genes (pdu) of Salmonella typhimurium: three
RT genes for the propanediol dehydratase.";
RL J. Bacteriol. 179:6633-6639(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=S.enterica subsp. enterica serovar Typhimurium; STRAIN=LT2;
RX MEDLINE=94429843; PubMed=10498708;
RA Bobik T.A., Havemann G.D., Busch R.J., Williams D.S., Aldrich H.C.;
RT "The propanediol utilization (pdu) operon of salmonella enterica
RT serovar typhimurium lt2 includes genes necessary for formation of
RT polyhedral organellae involved in coenzyme B(12)-dependent 1, 2-
RT propanediol degradation.";
RL J. Bacteriol. 181:5967-5975(1999).
DR EMBL; AE008790; AAL20956.1; -.
DR EMBL; AF026270; AAD39016.1; -.
DR InterPro; IPR001670; Fe-ADH.
DR Pfam; PF00465; Fe-ADH; 1.
DR PROSITE; PS00913; ADH_IRON_1; 1.
KW Complete proteome.
SQ SEQUENCE 370 AA; 39469 MW; 90BFD08FEB08B475 CRC64;

Query Match          64.3%; Score 36; DB 16; Length 370;
Best Local Similarity 63.6%; Pred. No. 1.9e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALVQOMEQLRQ 12
DB 305 ALIQOIELLKQ 315

RESULT 35
Q9VAY9 PRELIMINARY; PRT; 380 AA.
AC Q9VAY9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE CG8269 protein (LD07994P).
GN DMN OR CG8269.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.

```

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

DT	01-OCT-2000 (TReMBLrel. 15, last sequence update)
DT	01-OCT-2002 (TReMBLrel. 22, last annotation update)
DE	Hypothetical protein BHI1848.
GN	BHI1848.
OS	Bacillus halodurans.
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX	NCBI_TaxID=86665;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C-125 / JCM 9153;
RX	MEDLINE=20512582; PubMed=11058132;
RA	Takami H., Nakaseko K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA	Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA	Horikoshi K.;
RT	"Complete genome sequence of the alkaliphilic bacterium Bacillus
RT	halodurans and genomic sequence comparison with Bacillus subtilis.";
RL	Nucleic Acids Res. 28:3317-4331(2000).
DR	EMBL: AP001513; BAB05567.1; -
DR	InterPro: IPR000759; Adnrx_reductase.
DR	InterPro: IPR002937; Amino oxidase.
DR	InterPro: IPR000171; Bac_phycoene_dh.
DR	InterPro: IPR000205; NAD binding.
DR	Pfam: PF01593; Amino oxidase; 1.
DR	PRINTS: PR00419; ADXRDPTASE.
DR	ProDom: PD139017; Bac_phycoene_dh; 1.
DR	Hypothetical protein; Complete proteome.
SW	SEQUENCE 498 AA; 55939 MW; 59E5542E381CD7C CRC64;
QY	Query Match
Best Local Similarity	64.3%; Score 36; DB 16; Length 498;
Matches	7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db	130 KALVOEMEXL 139
QY	1 KALVOEMEXL 10
Matches	7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db	130 KALVOEMEXL 139
RESULT 37	
Q8MSAO	PRELIMINARY; PRT; 521 AA.
ID	Q8MSAO
AC	Q8MSAO
DT	01-OCT-2002 (TReMBLrel. 22, Created)
DT	01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT	01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE	LP02637P.
GN	CG4465.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Berkelley;
RC	Stapleton M., Brokstein P., Hong L., Ashyani A., Carlson J.,
RA	Champe M., Chavez C., Dorsett V., Dresnak D., Farfan D., Frise E.,
RA	Guerra R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA	Miranda A., Mungall C.J., Nuno J., Paoletti J., Paragas V., Park S.,
RA	Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA	Celniker S.;
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL: AY118989; AAM50849.1; -
DR	FLYBase: FBgn0038750; CG4465.
SW	SEQUENCE 521 AA; 57981 MW; D63716D7D77C55D9 CRC64;
QY	Query Match
Best Local Similarity	64.3%; Score 36; DB 5; Length 521;
Matches	7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Db	1 KALVOEMEXL 12
Matches	7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Db	287 KALVOEMEXL 298

```

RESULT 38
Q9BWC1 PRELIMINARY: PRT: 530 AA.
ID Q9BWC1
AC Q9BWC1
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Similar to pericentriolar material 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Homidae; Homo.
OC NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue:lung;
RA Strauberg R.;
RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL, BC000453, AA00453.1, -.
SQ SEQUENCE 530 AA; 59198 MW; 52B76CDE001E0690 CRC64;

```

```

Query Match 64.3%; Score 36; DB 4; Length 530;
Best Local Similarity 63.6%; Pred. No. 2.7e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 KALVOQMEQLR 11
Db 323 KRLMQOEQLR 333

```

## RESULT 39

```

Q9VDS2 PRELIMINARY: PRT: 577 AA.
ID Q9VDS2
AC Q9VDS2
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE CG4465 protein.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Suton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abitl J.F., Abmayr A., An H.-J., Andrews-Pfannkoch C., Balwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabros B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Jaisin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
RA Jaisin D., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulip D., Lai Z.,
RA Laake P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy W., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

```

```

RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skipski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
DR EMBL, AE003728; AAF55718.1, -.
DR FlyBase; FBgn0038750; CG4465.
SQ SEQUENCE 577 AA; 64509 MW; 3BED796F8EDCF38 CRC64;

```

```

Query Match 64.3%; Score 36; DB 5; Length 577;
Best Local Similarity 58.3%; Pred. No. 2.9e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 KALVOQMEQLR 12
Db 343 KRLPQLELR 354

```

## RESULT 40

```

Q9H7J2 PRELIMINARY: PRT: 674 AA.
ID Q9H7J2
AC Q9H7J2
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE FLJ00087 protein (Fragment).
GN FLJ00087.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Homidae; Homo.
OC NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Ohara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen."
RT Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL, AK024498; BAB15778.1, -.
DR InterPro; IPR000008; RaGAP.
DR InterPro; IPR001936; RaGAP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00616; RaGAP; 1.
DR SMART; SM00323; RaGAP; 1.
DR PROSITE; PS50018; RAS_CTPASE_ACTIV_2; 1.
FT NON-TER
SQ SEQUENCE 674 AA; 74598 MW; D2621C9F32806553 CRC64;

```

```

Query Match 64.3%; Score 36; DB 4; Length 674;
Best Local Similarity 63.6%; Pred. No. 3.4e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 KALVOQMEQLR 11
Db 651 RALPQOEQLR 661

```

```

Search completed: September 8, 2003, 16:06:12
Job time : 36 secs

```